





; APPLICATION NUMBER: US/09/045,973  
 ; FILING DATE: Filed Herewith  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:

QY 340 tcatgacgcgtcacctgacttgcgcgggagatgccctgcgaacacgctgcgtgcgggagat 395  
||||| | | | | | | | | | | | | | | |  
pb 458 tcatgatctatcgagaacactgcgtctcgtgtagagagcattccagaacgctgcgaaggccacg - 515





```

STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,881
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BILLINGS, LUCY J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0470 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:

INFORMATION FOR SEQ. ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1691 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: MUSCNOT02
CLONE: 971204
US-09-013-881-12

```

Query Match	4.28;	Score	36.8;	DB	3;	Length	1691;
Best Local Similarity	54.58;	Pred.	No. 0.16;				
Matches	96;	Conservative	0;	Mismatches	77;	Indels	3;
						Gaps	1

Oy	273	ggtgaagagctctgtcgttgaactgctgcgggggtctcccaagagcgtgacacgtgtatc	332
Db	832	GGAGGGGAAGATCTCGTGTGCATTGGCTGTGGGCGCTGAGCCGATCCGCCACCTGTGTACTG	891
Oy	333	gcatacatcatagaccgttcaactgacttggctcggagagatgacctgaacacgtgtgc	392
Db	892	GCTTACCTCATGCTGTGTACACACACCTTACCCTCTGTGAGGGGCATCAAGAAAGTCAAGAC	951
Oy	333	gggagatctctgtgcacaaccacaagtggtcgttcagaagacaagctccaagagatttga	448
Db	952	CACCGA---GGCATTCATCCCAACCGGGGCTTCTCTGAGGCAAGCTCTGGGCCCTGGA	1004

RESULT 11  
 US-08-649-046-3  
 Application 3  
 US/08649046

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (496)..(498)
? OTHER_INFORMATION: Xaa = any amino acid
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (589)..(591)
? OTHER_INFORMATION: Xaa = any amino acid
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (592)..(594)
? OTHER_INFORMATION: Xaa = any amino acid
US-09-164-193-9

Query Match      3.9%; Score 34; DB 4; Length 627;
Best Local Similarity 53.8%; Pred. No. 0.67;
Matches 70; Conservative 0; Mismatches 60; Indels 0; Gaps 0

```

		3.9%	Score 34;	DB 4;	Length 627;
	Best Local Similarity	53.8%;	Pred.	No. 0.67;	
Matches	70;	Conservative	0;	Mismatches	60; Indels 0; Gaps 0;
QY	232	tcaagaagaatlaataatcattccacagagtgcggctcggtgagagcttcgttga	291		
Dd	350	tccaagaaggagtcacaatttgtctccaagtaccagtcgtggccaagtgtttaacgtgc	409		
OY	292	aactgcttgcccgggtctccagaagcgtgaacctgtgatcgatacatcatgacgtcaa	351		

[illegible]; Sequence 7, *P*

```

? APPLICANT: Acton, Susan L.
? TITLE OF INVENTION: NOVEL CSAPP NUCLEIC ACID MOLECULES AND USES THEREFOR (as ame
? FILE REFERENCE: NMT-051
? CURRENT APPLICATION NUMBER: US/09/164,193C
? CURRENT FILING DATE: 1998-09-30
? NUMBER OF SEQ ID NOS: 23
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 7
? LENGTH: 814
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(627)
? FEATURE:

```

```

OTHER INFORMATION: 2,12,496-498,786,803,804,806,808,810 and 811 may
OTHER INFORMATION: be any nucleotide
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(3)
OTHER INFORMATION: Xaa = any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10)..(12)
OTHER INFORMATION: Xaa = any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13)..(15)
OTHER INFORMATION: Xaa = any amino acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: (496)..(498)
OTHER INFORMATION: Xaa = any amino acid
FEATURE:
NAME/KEY: misc_feature

```

CLONE: PTZgpt-F1s

TOPOLOGY: linear



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 16:09:41 ; Search time 23.37 Seconds

(without alignments)  
529.321 Million cell updates/sec

Title: US-09-544-525-2

Perfect score: 897  
Sequence: 1 MCGMKNKILPGIYGNFKDA.....AKNIIAPEILKFWAFRLRL 167

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_1101:\*

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:\*

2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:\*

3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:\*

4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:\*

5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:\*

6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:\*

7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:\*

8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:\*

9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:\*

10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:\*

11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:\*

12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:\*

13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:\*

14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:\*

15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:\*

16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:\*

17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:\*

18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:\*

19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:\*

20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:\*

21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:\*

22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	891	99.3	167	21	AA823298 Human dual-specific
2	831.5	92.7	184	21	AA823298 Human dual-specific
3	831.5	92.7	184	22	AA823298 Human dual-specific
4	831.5	92.7	184	22	AA823298 Human dual-specific
5	831.5	92.7	184	22	AA823298 Human dual-specific
6	791.5	88.2	184	22	AA823298 Human dual-specific
7	753.5	84.0	205	22	AA823298 Human dual-specific
8	715.5	79.8	205	22	AA823298 Human dual-specific
9	702.5	78.3	205	22	AA823298 Human dual-specific
10	644	71.8	145	22	AA823298 Human dual-specific
11	458.5	51.1	329	22	AA823298 Human dual-specific

12	433.5	48.3	96	21	AA823298 Human dual-specific
13	210	23.4	482	21	AA823298 Human dual-specific
14	210	23.4	482	21	AA823298 Human dual-specific
15	210	23.4	482	21	AA823298 Human dual-specific
16	209	23.3	218	22	AA823298 Human dual-specific
17	207	23.1	179	22	AA823298 Human dual-specific
18	207	23.1	341	22	AA823298 Human dual-specific
19	207	23.1	659	22	AA823298 Human dual-specific
20	205	22.9	157	21	AA823298 Human dual-specific
21	205	22.9	157	21	AA823298 Human dual-specific
22	203.5	22.7	170	22	AA823298 Human dual-specific
23	203.5	22.7	170	22	AA823298 Human dual-specific
24	203.5	22.7	663	18	AA823298 Human dual-specific
25	197	22.0	168	22	AA823298 Human dual-specific
26	197	22.0	168	22	AA823298 Human dual-specific
27	196	21.9	170	22	AA823298 Human dual-specific
28	196	21.9	170	22	AA823298 Human dual-specific
29	196	21.9	189	21	AA823298 Human dual-specific
30	191	21.3	170	22	AA823298 Human dual-specific
31	191	21.3	552	22	AA823298 Human dual-specific
32	191	21.3	552	22	AA823298 Human dual-specific
33	191	21.3	1049	22	AA823298 Human dual-specific
34	187	20.8	661	21	AA823298 Human dual-specific
35	186.5	20.8	665	22	AA823298 Human dual-specific
36	186.5	20.8	666	22	AA823298 Human dual-specific
37	186.5	20.8	672	22	AA823298 Human dual-specific
38	186	20.7	171	22	AA823298 Human dual-specific
39	186	20.7	171	22	AA823298 Human dual-specific
40	185	20.6	221	21	AA823298 Human dual-specific
41	183.5	20.5	223	22	AA823298 Human dual-specific
42	183.5	20.5	236	22	AA823298 Human dual-specific
43	181	20.2	552	22	AA823298 Human dual-specific
44	175	19.5	220	22	AA823298 Human dual-specific
45	175	19.5	784	21	AA823298 Human dual-specific

#### ALIGNMENTS

RESULT 1	
AA823298	standard; Protein: 167 AA.
AA823298:	
09-FEB-2001 (first entry)	
Human dual-specificity MAP kinase phosphatase DSP-3.	
Human DSP-3; dual-specificity phosphatase; antibody;	
dual-specificity MAP kinase phosphatase family; PTP; recombinant production;	
protein tyrosine phosphatase family; PTP; recombinant production;	
proliferative response; cell differentiation; cell survival;	
proliferative disorder; cell cycle abnormality; metabolic disease;	
Duchenne muscular dystrophy; cancer; graft-versus-host disease;	
autoimmune disease; allergy.	
Homo sapiens.	
OS	
XX	
PN	W0200060092-A2.
XX	
PD	12-OCT-2000.
XX	
PP	07-APR-2000; 2000WO-US09185.
XX	
PR	07-APR-1999; 99US-0128225.
PR	02-JUL-1999; 99US-0142338.
XX	
PA	(CEPT-) CEPTYR INC.
XX	
PI	Lucho RM, Wei B;
XX	
DR	WPI: 2000-665011/64.



```
RESULT 3
AAB73216
ID AAB73216 standard; Protein; 184 AA.
XX
XX AAB73216;
XX
XX 11-MAY-2001 (first entry)
XX
DE Human phosphatase AAB74753_h.
XX
XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
XX cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
XX congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
XX Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
XX schizophrenia; hamartoma.
XX
XX Homo sapiens.
XX
XX WO200112819-A2.
XX
XX 22-FEB-2001.
XX
XX 11-AUG-2000; 2000WO-US22158.
XX
XX 13-AUG-1999; 99US-0149005.
XX
XX (SUGEN-) SUGEN INC.
XX
XX PLOWMAN GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
XX WPI; 2001-211226/21.
XX
XX N-PSDB; AAF63568.
XX
XX
XX New protein phosphatase polypeptide for diagnosing and treating
XX phosphatase related disorders such as cancer, schizophrenia, cardiac
XX dysfunction and/or vascular disorders
XX
XX
XX Claim 6; Fig 5; 138pp; English.
XX
XX The present invention relates to phosphatase proteins and coding
XX sequences. The present sequence is one such phosphatase. Phosphatases are
XX enzymes that catalyse the dephosphorylation of proteins modified by
XX phosphorylation of serine, threonine or tyrosine residues. The
XX phosphatases are useful for treating a variety of diseases: for example
XX cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,
XX synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,
XX hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
XX glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
XX hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
XX congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
XX ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
XX Zonana syndrome, schizophrenia and hamartomas.
XX
XX Sequence 184 AA;
XX
XX
XX Query Match 92.7%; Score 831.5; DB 22; Length 184;
XX Best Local Similarity 88.6%; Pred. No. 9e-94;
XX Matches 163; Conservative 0; Mismatches 4; Indels 17; Gaps 2;
```

```
RESULT 4
AAB67167
ID AAB67167 standard; Protein; 184 AA.
XX
XX AAB67167;
XX
XX 12-APR-2001 (first entry)
XX
XX Human dual-specificity phosphatase DSP-3.
XX
XX DSP-3; dual-specificity phosphatase; cell proliferation;
XX cell signalling; cancer; graft-versus-host disease; autoimmune disease;
XX allergy; metabolic disease; Duchenne muscular dystrophy.
XX
XX Homo sapiens.
XX
XX WO200102581-A1.
XX
XX 11-JAN-2001.
XX
XX 20-APR-2000; 2000WO-US10868.
XX
XX 02-JUL-1999; 99US-0142338.
XX
XX 07-APR-2000; 2000WO-US09185.
XX
XX (CEPT-) CEPTYR-INC.
XX
XX LUCHE RM, Wei B;
XX WPI; 2001-138148/14.
XX
XX N-PSDB; AAF32191.
XX
XX
XX New dual-specificity phosphatase-3 polypeptide and its variants useful
XX for treating disorders associated with DSP-3 activity, defects in cell
XX proliferation, differentiation or survival, e.g. Duchenne muscular
XX dystrophy, cancer
XX
XX
XX Claim 1; Fig 2; 70pp; English.
XX
XX The present invention provides the protein and coding sequences of the
XX human dual-specificity phosphatase DSP-3. The DSP-3 protein is involved
XX in cell signalling and the sequences can be used in the treatment of
XX cancer, metabolic and autoimmune diseases, allergies, graft-versus-host
XX disease, abnormal cell proliferation and Duchenne muscular dystrophy.
XX
XX Sequence 184 AA;
XX
XX
XX Query Match 92.7%; Score 831.5; DB 22; Length 184;
XX Best Local Similarity 88.6%; Pred. No. 9e-94;
XX Matches 163; Conservative 0; Mismatches 4; Indels 17; Gaps 2;
```

RESULT 5  
AAB66431  
ID AAB66431 standard; Protein; 184 AA.  
XX  
AC AAB66431;  
XX  
DT 06-APR-2001 (first entry)  
XX  
DE Human DSP-3 protein.  
XX  
KW Human; DSP-3; cytosolic; immunosuppressive; anti-allergic;  
KW dual specificity phosphatase-3; cell proliferation; metabolic diseases;  
KW Duchenne muscular dystrophy; cancer; graft-versus-host disease;  
KW autoimmune disease; allergy.  
XX  
OS Homo sapiens.  
XX  
PN WO200102582-A1.  
XX  
PD 11-JAN-2001.  
XX  
PE 29-JUN-2000; 2000WO-US18207.  
XX  
PR 02-JUL-1999; 99US-0142338.  
PR 07-APR-2000; 2000WO-US09185.  
PR 20-APR-2000; 2000WO-US10868.  
XX  
PA (CEPT-) CEPTYR INC.  
XX  
PI Luche RM, Wei B;  
XX  
DR WPI: 2001-138149/14.  
DR N-PSDB; AAF29601.  
XX  
PT New dual-specificity phosphatase (DSP)-3 and DSP-3 alternate form  
PT polypeptides, useful for identifying modulators DSP-3 or DSP-3  
PT alternate form activity, especially for treating e.g. cancer,  
PT autoimmune diseases or allergies -  
XX  
PS Claim 1; Fig 2; 86pp; English.  
XX  
CC The present sequence is given in a specification providing human dual  
CC specificity phosphatase-3 (DSP-3) and a murine DSP-3 variant polypeptide.  
CC The polypeptides are useful for dephosphorylating a substrate of DSP-3.  
CC e.g. MAP-kinase. They may be used to treat or prevent diseases associated  
CC with cell proliferation, immunosuppression, metabolic diseases, or  
CC abnormal cell growth or cell cycle abnormalities. They are also useful  
CC for identifying agents that modulate their activity. The modulators are  
CC useful for treating disorders associated with DSP-3 or DSP-3 variant  
CC activity, e.g. Duchenne muscular dystrophy, cancer, graft-versus-host  
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal  
CC cell growth, abnormal cell proliferation and cell cycle abnormalities.  
CC The modulating agents are useful for modulating, modifying or altering  
CC cellular responses, e.g. in vivo or in vitro cell proliferation,  
CC differentiation or survival.  
XX  
SO Sequence 184 AA;  
SO  
Query Match 92.7%; Score 831.5; DB 22; Length 184;  
Best Local Similarity 88.6%; Pred. NO. 9e-94;  
Matches 163; Conservative 0; Mismatches 4; Indels 17; Gaps 2;  
OY 1 MGNQNKILPLGLYIGNFKDARDAEQLSKNKVTHILSVHDSFG-----LCW----- 45  
Db 1 mngnmkllppllygnfkardaeqlsknkvtllsvndarpmlegvkylcipadps 60  
OY 46 --RRHRESIKFTHECHLRESCVHCLAGVSRVTLVIYIMVDPFGMEDALHTVRA 103  
Db 61 qnltrhfkesikflheclrgescvlhclagvsrvtlviyimvdpfgmedalhtvra 120  
OY 104 GRSCANPVWGFORLOEFKEHEVHQYRQWLKEEYGESPLQDAEAKNLIAPGILKFWAF 163  
|||||

Db 121 grscanpvwfgfgrqlqefekhevhqyrqwlkeeygesplqdaeeaknliapgilkwaf 180  
OY 164 LRRL 167  
Db 181 lrri 184  
RESULT 6  
AAB73213  
ID AAB73213 standard; Protein; 184 AA.  
XX  
AC AAB73213;  
XX  
DT 11-MAY-2001 (first entry)  
XX  
DE Murine phosphatase AA103595\_m.  
XX  
KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
KW congenital muscle disorder; Papillon-LeFevre syndrome; Cowden disease;  
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
KW schizophrenia; hamartoma.  
XX  
OS Mus sp.  
XX  
PN WO200112819-A2.  
XX  
PD 22-FEB-2001.  
XX  
PE 11-AUG-2000; 2000WO-US22158.  
XX  
PR 13-AUG-1999; 99US-0149005.  
XX  
PA (SUGEN-) SUGEN INC.  
XX  
PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;  
XX  
DR WPI: 20014211226/21.  
DR N-PSDB; AAF63565.  
XX  
PT New protein phosphatase polypeptide for diagnosing and treating  
PT phosphatase related disorders such as cancer, schizophrenia, cardiac  
PT dysfunction and/or vascular disorders -  
XX  
PS Claim 6; Fig 5; 138pp; English.  
XX  
CC The present invention relates to phosphatase proteins and coding  
CC sequences. The present sequence is one such phosphatase. Phosphatases are  
CC enzymes that catalyse the dephosphorylation of proteins modified by  
CC phosphorylation of serine, threonine or tyrosine residues. The  
CC phosphatases are useful for treating a variety of diseases: for example  
CC cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,  
CC synovial sarcomas, renal cell carcinoma, non-small cell lung cancer,  
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,  
CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological  
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,  
CC congenital muscle disorders, Papillon-LeFevre syndrome, Cowden disease,  
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan  
CC Zonana syndrome, schizophrenia and hamartomas.  
XX  
SO Sequence 184 AA;  
SO  
Query Match 88.2%; Score 791.5; DB 22; Length 184;  
Best Local Similarity 82.6%; Pred. NO. 7.3e-89;  
Matches 152; Conservative 11; Mismatches 4; Indels 17; Gaps 2;  
OY 1 MGNQNKILPLGLYIGNFKDARDAEQLSKNKVTHILSVHDSFG-----LCW----- 45  
Db 1 mngsmqllppllygnfkardaeqlsrnkvtllsvndarpmlegvkylcipadtps 60  
OY 46 --RRHRESIKFTHECHLRESCVHCLAGVSRVTLVIYIMVDPFGMEDALHTVRA 103  
|||||



Query Match	84.0%; Score 753.5; DB 22; Length 205;
Best Local Similarity	79.5%; Pred. No. 3.9e-84;
Matches 151; Conservative 3; Mismatches 9; Indels 27; Gaps 3	
QY 1 MGNGNKKILPLGLYIGNFQKARDADAEOLSKNKVTHILSVHDSFG-----LCW----- 45	
Pb 1 mgnmgnkllplglyignfdaeaqelskknkvchllsvhsaspmlegvkyllcpaadsp 60	
QY 46 --RRRHEFESIKFTHIECRIRGESCIVHCLAGYSRSYTLVIATIMYTDRCWEDALFTVRA 103	
Pb 61 qnltrhfkesikfthiecrirgescivhclagysrsylvliaylmtvtdfgwedaltlvra 120	
QY 104 GRSCNPNWVQPCROLOEFPEKHEVHQYRQWLKEEYGSPIQDAEFAKNILA----- 153	
Pb 121 grscnnpwvqtrqqlqefekhevhqyrqwlkeeygespiqdaeeaknllqkykegrtep 180	
QY 154 APGILKFWAF 163	
Pb 181 qpgarrwssf 190	
RESULT 8	
AAB66443	
ID AAB66443 standard; Protein; 205 AA.	
AC AAB66443;	
XX	
DT 06-APR-2001 (first entry)	
XX	
DE Murine DSP-3 variant.	
XX	
KW Mouse; DSP-3; cytosolic; immunosuppressive; antiallergic;	
KM dual specificity phosphatase-3; cell proliferation; metabolic diseases;	
KW Duchenne muscular dystrophy; cancer; graft-versus-host disease;	
XX autoimmune disease; allergy; variant.	
XX	
OS Mus musculus.	
XX	
FN W0200102582-A1.	
XX	
PD 11-JAN-2001.	
XX	
PF 29-JUN-2000; 2000WO-US18207.	
XX	
PR 02-JUL-1999; 99US-0142338.	
PR 07-APR-2000; 2000WO-US09185.	
PR 20-APR-2000; 2000WO-US10868.	
XX	
PA (CEPT-) CEPTYR INC.	
XX	
PI Lucbe RM, Wei B;	
XX	
DR WPI: 2001-138149/14.	
DR N-PSDB; AAF29608.	
XX	
XX New dual-specificity phosphatase (DSP)-3 and DSP-3 alternate form	
PT polypeptides, useful for identifying modulators DSP-3 or DSP-3	
PT alternate form activity, especially for treating e.g. cancer,	
PT autoimmune diseases or allergies -	
XX	
PS Claim 50; Fig 6; 86pp; English.	
XX	
CC The present sequence is given in a specification providing human dual	
CC specificity phosphatase-3 (DSP-3) and a murine DSP-3 variant polypeptide.	
CC The polypeptides are useful for dephosphorylating a substrate of DSP-3,	
CC e.g. MAP-kinase. They may be used to treat or prevent diseases associated	
CC with cell proliferation, immunosuppression, metabolic diseases, or	
CC abnormal cell growth or cell cycle abnormalities. They are also useful	

CC for identifying agents that modulate their activity. The modulators are  
 CC useful for treating disorders associated with DSP-3 or DSP-3 variant  
 CC activity, e.g. Duchenne muscular dystrophy, cancer, graft-versus-host  
 CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal  
 CC cell growth, abnormal cell proliferation and cell cycle abnormalities.  
 CC The modulating agents are useful for modulating, modifying or altering  
 CC cellular responses, e.g. in vivo or in vitro cell proliferation,  
 CC differentiation or survival.

XX Sequence 205 AA:

Query Match 79.8%; Score 715.5; DB 22; Length 205;  
 Best Local Similarity 81.7%; Pred. No. 1.8e-79;  
 Matches 138; Conservative 10; Mismatches 4; Indels 17; Gaps 2;

QY 1 MGNGMKNKILPGLYIGNFKARDAEQSKKNVTHILSVHDSFG-----LCW----- 45  
 DB 1 msgsmqqlpglylgnfkardaeqlsrnkvhllsvhdtarpmlegvlylcipaadtps 60  
 QY 46 --RTRHFKESIKFIHECRLRGESCLVHCLAGYSRSVTLVIAYIMVTDFGWDALHTVRA 103  
 DB 61 gnltrhfkesikfihecrilqgesclvhclagysrsvtlviayimvtdfgwdalhtvra 120

QY 104 GRSCANPNVGFQROLOEFKEHVOYROWLKEEYGESPLQDAEAKNLL 152  
 DB 121 grscanpnlgfgrqfqefekhevhyrqlreeygenplrddeeknll 169

RESULT 9  
 AAB81106  
 ID AAB81106 standard; Protein: 205 AA.

AC AAB81106;  
 DT 29-JUN-2001 (first entry)

DE Murine JNK activating phosphatase (JNAP).

KM C-Jun amino-terminal kinase; JNK; JNK activating phosphatase; JNAP;  
 KW dual-specificity phosphatase; ischemic heart disease; kidney failure;  
 KW cardiac hypertrophy; burn; liver damage; respiratory distress syndrome;  
 KW septic shock; rheumatoid arthritis; autoimmune disorder; mouse;  
 KW inflammatory disease; chromosome 13.

OS Mus sp.

PN WO200121812-A1.

PD 29-MAR-2001.

PF 21-SEP-2000; 2000WO-US25948.

PR 21-SEP-1999; 99US-0155068.

PR 20-SEP-2000; 2000US-0665819.

PA (AMGE-) AMGEN INC.

PA (BAYU) BAYLOR COLLEGE MEDICINE.

PI Tan T, Zhou G, Belmont JW, Fletcher FA, Chen AJ, Jurecic R;

DR WPI: 2001-316103/33.

DR N-PSDB: AAF86153.

PT New human or mouse c-Jun amino-terminal kinase (JNK) activating  
 PT phosphatase is useful for diagnosing or treating diseases associated  
 PT with JNK activating phosphatase or JNK-mediated disorders, e.g.  
 PT inflammatory diseases

XX Claim 8; Fig 2; 116pp; English.

XX This invention relates to human or murine c-Jun amino-terminal kinase  
 CC (JNAP) activating phosphatase (JNAP). The present sequence represents

CC murine JNAP. The JNAP gene has been mapped to murine chromosome 13. The  
 CC JNAP protein has an amino-terminal dual-specificity phosphatase domain  
 CC and lacks a non-catalytic carboxy-terminal domain. The protein has  
 CC functional activities comprising dual-specificity phosphatase activity  
 CC and JNK pathway upregulation activity. The JNAP polypeptide can be used  
 CC for diagnosing a JNK-mediated disorder or a susceptibility to a  
 CC JNK-mediated disorder, and for identifying a compound which binds to it.  
 CC The JNAP polypeptide and nucleic acid molecules are useful in screening  
 CC for therapeutic agents to treat, prevent and/or detect conditions  
 CC relating to JNK-mediated disorders, e.g. ischemic heart disease, cardiac  
 CC hypertrophy, burns due to heat or radiation, kidney failure, liver damage  
 CC due to oxidative stress or alcohol, respiratory distress syndrome, septic  
 CC shock, rheumatoid arthritis, autoimmune disorders, or other types of  
 CC inflammatory diseases. The polypeptides may also be used for screening  
 CC therapeutic agents for treating a disease involving cytokine production  
 CC in an animal, and for identifying agents that modulate the JNK signal  
 CC transduction pathways. The nucleic acid molecules encoding JNAP are  
 CC useful for mapping the locations of the JNK activating phosphatase gene  
 CC and related genes on chromosomes; as anti-sense inhibitors of JNK  
 CC activating phosphatase polypeptide expression; or for gene therapy.

XX Sequence 205 AA:

Query Match 78.3%; Score 702.5; DB 22; Length 205;  
 Best Local Similarity 80.5%; Pred. No. 7.1e-78;  
 Matches 136; Conservative 10; Mismatches 6; Indels 17; Gaps 2;

QY 1 MGNGMKNKILPGLYIGNFKARDAEQSKKNVTHILSVHDSFG-----LCW----- 45  
 DB 1 msgsmqqlpglylgnfkardaeqlsrnkvhllsvhdtarpmlegvlylcipaadtps 60  
 QY 46 --RTRHFKESIKFIHECRLRGESCLVHCLAGYSRSVTLVIAYIMVTDFGWDALHTVRA 103  
 DB 61 gnltrhfkesikfihecrilqgesclvhclagysrsvtlviayimvtdfgwdalhtvra 120

QY 104 GRSCANPNVGFQROLOEFKEHVOYROWLKEEYGESPLQDAEAKNLL 152  
 DB 121 grscanpnlgfgrqfqefekhevhyrqlreeygenplrddeeknll 169

RESULT 10  
 AAB66442  
 ID AAB66442 standard; Protein: 145 AA.

AC AAB66442;

DT 06-APR-2001 (first entry)

DE Human MAP-kinase phosphatase DSP-3.

KM Human; MAP-kinase; mitogen-activated protein kinase; DSP-3; cytosolic;  
 KW immunosuppressive; antiallergic; dual specificity phosphatase-3;  
 KW cell proliferation; metabolic diseases; Duchenne muscular dystrophy;  
 KW cancer; graft-versus-host disease; autoimmune disease; allergy; DSP-3.

OS Homo sapiens.

PN WO200102582-A1.

PD 11-JAN-2001.

PF 29-JUN-2000; 2000WO-US18207.

PR 02-JUL-1999; 99US-0142338.

PR 07-APR-2000; 2000WO-US09185.

PR 20-APR-2000; 2000WO-US10868.

PA (CEPT-) CEPTAR INC.

PI Luche RM, Wei B;

DR WPI: 2001-138149/14.

```

XX New dual-specificity phosphatase (DSP)-3 and DSP-3 alternate form
PT polypeptides, useful for identifying modulators DSP-3 or DSP-3
PT alternate form activity, especially for treating e.g. cancer,
PT autoimmune diseases or allergies
XX
XX Example 1; Fig 3; 86pp; English.
XX
CC The present sequence is given in a specification providing human dual
CC specificity phosphatase-3 (DSP-3) and a murine DSP-3 variant polypeptide.
CC The polypeptides are useful for dephosphorylating a substrate of DSP-3,
CC e.g. MAP-kinase. They may be used to treat or prevent diseases associated
CC with cell proliferation, immunosuppression, metabolic diseases, or
CC abnormal cell growth or cell cycle abnormalities. They are also useful
CC for identifying agents that modulate their activity. The modulators are
CC useful for treating disorders associated with DSP-3 or DSP-3 variant
CC activity, e.g. Duchenne muscular dystrophy, cancer, graft-versus-host
CC disease, autoimmune diseases, allergies, metabolic diseases, abnormal
CC cell growth, abnormal cell proliferation and cell cycle abnormalities.
CC The modulating agents are useful for modulating, modifying or altering
CC cellular responses, e.g. in vivo or in vitro cell proliferation,
CC differentiation or survival.
CC
XX
XX Sequence 145 AA:
SO
Query Match 71.8%; Score 644; DB 22; Length 145;
Best Local Similarity 87.6%; Pred. No. 6.6e-71;
Matches 127; Conservative 0; Mismatches 4; Indels 14; Gaps 2;
OY 1 MGNGMKNKILPGIYIGNFKDARDADQLSKNKVTHILSVHDSFG-----ICW-----R 46
Db 1 mgngmknkllpglyignfkardaeqlsknkvtllsvhdsarpmlegvkylicipadps 60
OY 47 TRHFKESIKFHECRRLRGESCLVHCAGVSRSVTLVYIMVTYDFGEMDLHTYRAGRS 106
Db 61 trhkessikfhecrklrgescvchlagvsrsvtlvayimvtidfgwedahlvtrags 120
OY 107 CANPNVGFQROLOEFKEHEVHQYRQ 131
Db 121 canpnvgfqrqlqefekhevhqyrq 145
RESULT 11
AAE04840
AAE04840 standard; Protein: 329 AA.
XX
AC AAE04840;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human SGP008 phosphatase polypeptide.
XX
KW Human; SGP008 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cystostasis;
KW neurological disorder; vinclozide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnary; tranquilliser; antisthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; mitogen; chromosome 20q11.2.
XX
OS Homo sapiens.
XX
FH Key
FT 1..173 Location/Qualifiers
FT Domain /label=Catalytic_domain
FT Domain 98..235
FT /label=Phosphatase_domain

```

```

FT Misc-difference 323..329
FT /note="Encoded by CTCGC"
XX
XX WO200146394-A2.
XX
XX 28-JUN-2001.
XX
XX 21-DEC-2000; 2000WO-US34736.
XX
XX 21-DEC-1999; 99US-0173255.
XX 28-DEC-1999; 99US-0175766.
XX 25-JAN-2000; 2000US-0178078.
XX 31-JAN-2000; 2000US-0179301.
XX
XX (SUGEN-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
XX Flanagan P;
XX
XX WPI: 2001-418058/44.
XX
XX N-PSDB; AAD09498.
XX
XX
XX Novel phosphatase polypeptide useful for treating cancers,
XX immune-related diseases and disorders, cardiovascular disease, brain or
XX neuronal associated diseases and metabolic disorders
XX
XX Claim 7; Fig 2; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
XX sequences encoding them, as well as various products and methods useful
XX for the diagnosis and treatment of various phosphatase-related diseases
XX and conditions. Substance that modulates the activity of phosphatase
XX polypeptide is used to treat immune-related diseases and disorders,
XX cardiovascular disease, brain or neuronal associated diseases and
XX metabolic disorders, including cancers of tissues, cancers of
XX haematopoietic origin, diseases of central and peripheral nervous
XX system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
XX amyotrophic lateral sclerosis, viral infections, infections caused by
XX prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
XX sexual dysfunction, mood disorders, attention disorders, cognition
XX disorders, hypotension, hypertension, psychotic disorders, neurological
XX disorders, dyskinesias and organ transplant rejection. The present
XX amino acid sequence is human SGP008 phosphatase polypeptide. This
XX sequence is classified as dual specificity phosphatase (DSP) and MAP
XX kinase phosphatase (MKP). SGP008 gene maps to chromosomal position
XX 20q11.2.
XX
XX Sequence 329 AA:
SO
Query Match 51.1%; Score 458.5; DB 22; Length 329;
Best Local Similarity 50.3%; Pred. No. 1.2e-47;
Matches 85; Conservative 27; Mismatches 40; Indels 17; Gaps 1;
OY 1 MGNGMKNKILPGIYIGNFKDARDADQLSKNKVTHILSVHDSFGICWR----- 46
Db 95 mgngmktkvipgylignfkardakldqgrnkthihsihesppllqdltylrtpadtpe 154
OY 47 ---TRHFKESIKFHECRRLRGESCLVHCAGVSRSVTLVYIMVTYDFGEMDLHTYR 103
Db 155 vlkhkfkexchffincrcingncvlhcfagrsstltvayimvtvgwvdrleakka 214
OY 104 GSCANPNVGFQROLOEFKEHEVHQYRQWLKREYSGSPIDDAFEAKNIL 152
Db 215 trlpianpvgfqrqlqefekhevhqyrqwlkreygsprideetrrall 263
RESULT 12
AAB41570
AAB41570 standard; Protein: 96 AA.
XX
AC AAB41570;
XX

```

DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF1334 polypeptide sequence SEQ ID NO:2668.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnerrary; antiparasitic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiallergic; immunosuppressant; cardiac;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;  
 KW antiviral; antibacterial; antifungal; antineumatic; antihypertensive;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI: 2000-602362/57.  
 DR N-PSDB: AAC75779.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11; Page 1911; 5507pp: English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerrary;  
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;  
 CC osteopathic; anticonvulsant; antiallergic; immunosuppressant;  
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antineoplastic; antibacterial; antiviral; antineumatic; antihypertensive;  
 CC antihypertensive; antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 CC  
 SO Sequence 96 AA;

Query Match 48.3%; Score 433.5; DB 21; Length 96;  
 Best Local Similarity 89.6%; Pred. NO. 2.5e-45;  
 Matches 86; Conservative 1; Mismatches 6; Indels 3; Gaps 2;

OY 1 MGNGANKILPGLYIGNFKARDADOLSKNKVTHILSVHDS--PGLCWRTRRHREKSTKFTH 58  
 DB 2 mgngmknkllpglylignfkardaeqlsknk-gnfvscprwpglclwrtrhfesikfih 60  
 OY 59 ECLRGESCIVHCLACVSRSVTLVIYVMTVDPEGM 94  
 DB 61 eclrlrgescivhclacvsvtlviyvmtdvdfgw 96  
 RESULT 13  
 AAB29641  
 ID AAB29641 standard; Protein; 482 AA.  
 AC  
 AC AAB29641;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human dual-specificity MAP kinase phosphatase DSP-10.  
 XX  
 KW Human DSP-10; dual-specificity phosphatase; antibody;  
 KW dual-specificity MAP kinase phosphatase; drug screening;  
 KW protein tyrosine phosphatase family; PTP; recombinant production;  
 KW proliferative response; cell differentiation; cell survival;  
 KW proliferative disorder; cell cycle abnormality; metabolic disease;  
 KW Duchenne muscular dystrophy; cancer; graft-versus-host disease;  
 KW autoimmune disease; allergy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200065068-A1.  
 XX  
 PD 02-NOV-2000.  
 XX  
 PF 20-APR-2000; 2000WO-US10966.  
 XX  
 PR 23-APR-1999; 99US-0130806.  
 XX  
 PA (CEPT-) CEPTYR INC.  
 XX  
 PI Luche RM, Wei B;  
 XX  
 DR WPI: 2000-687352/67.  
 DR N-PSDB: AAC64262.  
 XX  
 PT Novel dual specificity phosphatase-10 which dephosphorylates activated  
 PT mitogen-activated protein kinase, for identifying agents inhibiting  
 PT DSP-10 activity and modulate cell proliferation, differentiation,  
 PT survival -  
 XX  
 PS Claim 1; Fig 2; 65pp: English.  
 XX  
 CC The invention relates to a human dual-specificity mitogen-activated  
 CC protein (MAP) kinase phosphatase, DSP-10, and to nucleic acids encoding  
 CC it. The invention also relates to variants of DSP-10 which retain  
 CC activity, expression vectors and host cells comprising DSP-10 encoding  
 CC DNA, the recombinant production of DSP-10, an anti-DSP-10 antibody, and  
 CC a DSP-10 substrate-trapping mutant protein that has a reduced ability to  
 CC dephosphorylate a substrate relative to the wild-type DSP-10. The  
 CC invention additionally encompasses use of a DSP-10 modulator to modulate  
 CC a proliferative response, cell differentiation or cell survival. The  
 CC DSP-10 protein is useful for screening an agent that binds to DSP-10  
 CC and/or modulates DSP-10 activity, and is also useful for raising  
 CC antibodies. DNA encoding DSP-10 or a reporter protein is also useful for  
 CC screening an agent that modulates DSP-10 activity. The identified agents  
 CC that modulate DSP-10 activity are useful for treating Duchenne muscular  
 CC dystrophy, cancer, graft-versus-host disease, autoimmune diseases,  
 CC allergies, metabolic diseases, abnormal cell growth, abnormal cell  
 CC proliferation and cell cycle abnormalities. The DSP-10 modulating  
 CC agents inhibit or enhance signal transduction via a MAP-kinase  
 CC cascade, leading to cell proliferation. An agent that modulates DSP-10  
 CC activity may alter expression and/or stability of DSP-10. DSP-10  
 CC protein activity and/or the ability of DSP-10 to dephosphorylate a

CC substrate. Compounds so identified as capable of influencing DSP-10  
CC function (e.g., phosphorylation and/or phosphoserine/threonine  
CC dephosphorylation) are valuable for therapeutic and/or diagnostic  
CC purposes, since they permit treatment and/or detection of disease  
CC associated with DSP-10 activity. Such compounds are also valuable in  
CC research directed to molecular signaling mechanisms that involve  
CC DSP-10 and in the discovery and development of future DSP-10 compounds  
CC exhibiting greater specificity. The DSP-10 polypeptides are useful for  
CC dephosphorylating a DSP-10 substrate. DSP-10-specific antibodies  
CC and DSP-10 antisense probes are useful for detecting DSP-10 expression in  
CC a sample, and may thus be used in the detection of a proliferative  
CC disorder. The present sequence represents human DSP-10.

SO Sequence 482 AA;

Query Match 23.4%; Score 210; DB 21; Length 482;  
Best Local Similarity 36.0%; Pred. No. 6.8e-17;  
Matches 50; Conservative 21; Mismatches 48; Indels 20; Gaps 2;

QY 5 MNKILPGIYIGFKDARAEOLSK-----NKVTHILSVDSFGLCKRTR----- 48  
Db 322 ltlplpflfigneqdaqltmqrlnlyvlnvthlplhyekyglfnkrlpatdsnkq 381  
QY 49 -----HFKSIKFIHECRIRGESCIVHCLAGVSRTIVATIMVTDGMDALHTVRAG 104  
Db 382 nlrgyfeeaftleabqcgkqyllhbcqgvsatlylaimktrmtmdaykfkpk 441  
QY 105 RSCANPNVGFORLOEFK 123  
Db 442 rplspnlmfmgqllefee 460

RESULT 14

AA18655  
ID AAB18655 standard; Protein: 482 AA.

AC AAB18655;

DT 22-JAN-2001 (first entry)

XX A human regulator of intracellular phosphorylation.

KW Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;  
KW neurological disorder; Parkinson's disease; demyelinating disease;  
KW meningitis; developmental disorder; neuromuscular disorder; cancer;  
KW myasthenia gravis; cell proliferative disorder; actinic keratosis;  
KW arteriosclerosis; atherosclerosis; leukemia; melanoma; bronchitis;  
KW autoimmune disorder; inflammatory disorder; Addison's disease;  
KW acquired immunodeficiency disease; allergy; diabetes mellitus;  
KW rheumatoid arthritis; microbial infection; trauma.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 4 /note= "potential phosphorylation site"  
FT Modified-site 52 /note= "potential phosphorylation site"  
FT Modified-site 57 /note= "potential phosphorylation site"  
FT Modified-site 67 /note= "potential glycosylation site"  
FT Modified-site 140 /note= "potential phosphorylation site"  
FT Modified-site 140 /note= "potential phosphorylation site"  
FT Modified-site 146 /note= "potential phosphorylation site"  
FT Modified-site 160 /note= "potential phosphorylation site"  
FT Modified-site 184 /note= "potential phosphorylation site"  
FT Modified-site 201 /note= "potential glycosylation site"  
FT Modified-site /note= "potential phosphorylation site"

FT Modified-site 211 /note= "potential phosphorylation site"  
FT Modified-site 217 /note= "potential phosphorylation site"  
FT Modified-site 217 /note= "potential phosphorylation site"  
FT Modified-site 224 /note= "potential phosphorylation site"  
FT Modified-site 224 /note= "potential phosphorylation site"  
FT Modified-site 258 /note= "potential phosphorylation site"  
FT Modified-site 275 /note= "potential phosphorylation site"  
FT Modified-site 313 /note= "potential phosphorylation site"  
FT Modified-site 313 /note= "potential phosphorylation site"  
FT Active-site 321..461 /note= "potential phosphorylation site"  
FT Modified-site 353 /note= "dual-specificity phosphatase catalytic site"  
FT Modified-site 378 /note= "potential glycosylation site"  
FT Modified-site 430 /note= "potential phosphorylation site"  
FT Modified-site 467 /note= "potential phosphorylation site"  
FT Modified-site 472 /note= "potential phosphorylation site"  
FT Modified-site /note= "potential phosphorylation site"  
PN WO200055332-A2.  
PD 21-SEP-2000.  
PE 17-MAR-2000; 2000WO-US07277.  
PF 18-MAR-1999; 99US-0125593.  
PR 20-MAY-1999; 99US-0135049.  
PR 09-JUL-1999; 99US-0143188.  
XX (INCY-) INCYTE PHARM INC.  
XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
PI Lu DAM, Au-Young J;  
XX WPI: 2000-602121/57.  
DR N-PSDB: AAA75672.  
XX Novel human intracellular phosphorylation regulator polypeptides and  
PT polynucleotides for diagnosis, prevention and treatment of  
PT neurological, cell proliferative and autoimmune/inflammatory disorders  
PS Claim 1; Page 73-74; 96pp; English.  
XX The present sequence represents a human regulator of intracellular  
CC phosphorylation (HRIP). HRIP is useful for screening agonists and  
CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
CC are useful for treating a disease or condition associated with  
CC decreased or increased expression of functional HRIP. Diseases treated  
CC or diagnosed include neurological disorders such as stroke, Parkinson's  
CC disease, demyelinating diseases, bacterial and viral meningitis and  
CC other developmental disorders of the central nervous system,  
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
CC including leukemia, melanoma, myeloma and cancer of the adrenal gland,  
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
CC inflammatory disorder such as Addison's disease, acquired  
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
CC rheumatoid arthritis, microbial infection and trauma.  
SO Sequence 482 AA;

Query Match 23.4%; Score 210; DB 21; Length 482;  
Best Local Similarity 36.0%; Pred. No. 6.8e-17;

Matches 50; Conservative 21; Mismatches 48; Indels 20; Gaps 2;

QY 5 MNKILPGLYIGNFKDARDAEOLSK-----NKVTHILSVHDSPGICWRT----- 48  
 Db 322 ltpilpflfigneqdvrdidtmqrlniyvinvtthlplyhyekglfnykrilpatsnkg 381  
 QY 49 ----HKESIKFIHECRLAGESCLVHCLAGVSRSVTLVIAYIMTVTFEGMEDALHTVRAG 104  
 Db 382 nlrgyfeeaefeieeahqcgkylhhcgaqvsarsativaiymkhltmumtdaykfvkkg 441  
 QY 105 RSCANPNVGFQROLOEFK 123  
 Db 442 rplisplnfmqqllefee 460

## RESULT 15

AAB73225  
 ID AAB73225 standard; Protein; 482 AA.

AC AAB73225;

DT 11-MAY-2001 (first entry)

DE Human phosphatase MKP-5\_h.

KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
 KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
 KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;  
 KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
 KW schizophrenia; hamartoma.

OS Homo sapiens.

PN WO200112819-A2.

PD 22-FEB-2001.

PF 11-AUG-2000; 2000WO-US22158.

PR 13-AUG-1999; 99US-0149005.

PA (SUGC-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

DR WPI: 2001-211226/21.

DR N-PSDB: AAF63577.

XX New protein phosphatase polypeptide for diagnosing and treating  
 PT phosphatase related disorders such as cancer, schizophrenia, cardiac  
 PT dysfunction and/or vascular disorders -

PS Claim 6; Fig 5; 138pp; English.

XX The present invention relates to phosphatase proteins and coding  
 CC sequences. The present sequence is one such phosphatase. Phosphatases are  
 CC enzymes that catalyse the dephosphorylation of proteins modified by  
 CC phosphorylation of serine, threonine or tyrosine residues. The  
 CC phosphatases are useful for treating a variety of diseases: for example  
 CC cancer e.g. breast, urogenital, prostate, head, neck, lung cancers,  
 CC synovial sarcoma, renal cell carcinoma, non-small cell lung cancer,  
 CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,  
 CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological  
 CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,  
 CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,  
 CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan  
 CC Zonana syndrome, schizophrenia and hamartomas.

SO Sequence 482 AA;

Matches 50; Conservative 20; Mismatches 49; Indels 20; Gaps 2;

QY 5 MNKILPGLYIGNFKDARDAEOLSK-----NKVTHILSVHDSPGICWRT----- 48  
 Db 322 ltpilpflfigneqdvrdidtmqrlniyvinvtthlplyhyekglfnykrilpatsnkg 381  
 QY 49 ----HKESIKFIHECRLAGESCLVHCLAGVSRSVTLVIAYIMTVTFEGMEDALHTVRAG 104  
 Db 382 nlrgyfeeaefeieeahqcgkylhhcgaqvsarsativaiymkhltmumtdaykfvkkg 441  
 QY 105 RSCANPNVGFQROLOEFK 123  
 Db 442 rplisplnfmqqllefee 460

Search completed: February 12, 2002, 16:12:52  
 Job time: 191 sec

Query Match 23.4%; Score 210; DB 22; Length 482;  
 Best Local Similarity 36.0%; Pred. No. 6.8e-17;







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 12, 2002, 16:10:21 ; Search time 14.49 Seconds

(without alignments)  
877.926 Million cell updates/sec

Title: US-09-544-525-2

Perfect score: 897

Sequence: 1 MGNGMKNKILIPGLYIGNFKDA.....AKNIIAPGILKFWAFLRRL 167

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_68:\*\*

1: PIR1:\*\*

2: PIR2:\*\*

3: PIR3:\*\*

4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	208.5	23.2	619	2 T15969	hypothetical prote
2	202	22.5	330	2 T39698	protein tyrosine p
3	194	21.6	365	2 T32494	hypothetical prote
4	186	20.7	384	1 T38900	dual specificity p
5	185	20.6	276	2 T48906	protein-tyrosine-p
6	183.5	20.5	223	2 T49365	protein tyrosine p
7	182	20.3	142	2 T03074	dual specificity p
8	173.5	19.3	205	2 T49364	protein tyrosine p
9	170	19.0	364	1 S31304	protein-tyrosine-p
10	167.5	18.7	226	2 T21380	hypothetical prote
11	164	18.3	186	2 T16056	hypothetical prote
12	161	17.9	367	1 S29090	dual specificity p
13	159	17.7	303	2 T46405	hypothetical prote
14	159	17.7	393	2 A56947	dual specificity p
15	159	17.7	394	2 A56115	dual specificity p
16	158	17.6	314	1 A57126	dual specificity p
17	156	17.4	314	2 B57126	dual specificity p
18	152	16.9	367	1 S24411	dual specificity p
19	152	16.9	367	2 S52265	dual specificity p
20	147.5	16.4	185	1 A47196	dual specificity p
21	147	16.4	272	2 T18915	hypothetical prote
22	145	16.2	171	1 QOVZHI	dual specificity p
23	142	15.8	171	1 A42514	dual specificity p
24	139	15.5	171	1 T36845	dual specificity p
25	139	15.5	171	2 T28522	probable dual spec
26	139	15.5	171	2 B72161	JIL protein - vari
27	138.5	15.4	272	2 T19418	hypothetical prote
28	134	14.9	771	2 T47666	phosphatase-like p
29	132.5	14.8	283	2 G84438	probable protein p

30	131	14.6	171	1 B47452	dual specificity p
31	128	14.3	489	1 S58725	dual specificity p
32	120	13.4	209	1 S48459	probable dual spec
33	116.5	13.0	600	2 T18446	hypothetical prote
34	116	12.9	580	2 T18439	hypothetical prote
35	113	12.6	292	2 S41012	hypothetical prote
36	101.5	11.3	204	2 T17802	hypothetical prote
37	101	11.3	278	2 T39517	dual-specificity M
38	97	10.8	807	1 S44538	probable protein-t
39	95	10.6	160	2 T10278	protein tyrosine p
40	94	10.5	169	2 T30684	probable dual spec
41	92.5	10.3	597	1 S43743	probable dual spec
42	91.5	10.2	150	2 T21489	hypothetical prote
43	87	9.7	383	2 F84771	hypothetical prote
44	83	9.3	146	2 G71181	hypothetical prote
45	82	9.1	151	2 C75157	protein tyrosine p

## ALIGNMENTS

RESULT 1  
T15969  
Hypothetical protein F08B1.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T15969  
R:Chisoe, S.  
Submitted to the EMBL Data Library, July 1995  
A:Description: The sequence of C. elegans cosmid F08B1.  
A:Reference.number: Z18439  
A:Accession: T15969  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-619 <CH>  
A:Cross-references: EMBL:U23178; NID:9726421; PID:9726422; PIDN:AAC46719.1; CESP:F08  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:F08B1.1  
A:Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Query Match 23.2%; Score 208.5; DB 2; Length 619;  
Best Local Similarity 31.2%; Pred. No. 5.4e-13;  
Matches 53; Conservative 28; Mismatches 60; Indels 29; Gaps 4;

QY 2 GNGMKNKILIPGLYIGNFKDARDAEQLSKNKVTHI--LSVHDSGLGWR-----46  
DB 135 GDGILIPNIVLIGSQIDSLDETMDALDISVINLSMTCPSKVCIKEDKNFMRIPVND5 194  
QY 47 ----TRHFESIKRTHIECRLEGESCLVHCLAGVSRYTLVATYMTVTDEGMEDALATV 101  
DB 195 YQEKSLSPYFPMAYEFLEKCRRAKCKLIRCLAGISRSPTLAISYIMRYKMKSDAYRYV 254  
QY 102 RAGRCANPNVGFQROLOEFE---KHEVHQRQWIK-----EEGCEPL 142  
DB 255 KRRRSISPNFNMGLLEYENVLIKDHVLDVYNQASRPHRMHYGPSOL 304

RESULT 2  
T39698  
Protein tyrosine phosphatase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T39698

R:Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.  
Submitted to the EMBL Data Library, July 1999

A:Reference.number: Z21870

A:Accession: T39698

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-330 <WOO>

A:Cross-references: EMBL:AL109652; PIDN:CAB51765.1; GSPDB:GN000067

A: Experimental source: strain 97zh-; cosmid c17A3  
 C: Genetics:  
 A: Gene: p1041  
 A: Map position: 2

Query Match 22.5%; Score 202; DB 2; Length 330;  
 Best Local Similarity 32.3%; Pred. No. 1.2e-12;  
 Matches 52; Conservative 26; Mismatches 59; Indels 24; Gaps 4;

OY 3 NGNNKILPGLYIGNFKDARDAEQLSKNKYTHILSVHDS-SPGL-----CWR----- 46  
 DB 45 NDLSESKMLYLSMKTASELSTSDKGDYTLISAMSINPNLSPPEQOHLMTQIEDSSQ 104  
 OY 47 --TRHFKESIKFTHECRLGESCTVHCLAGVSRVTLVIATVTVDFGMDALHTVRAG 104  
 DB 105 NILQYREKKNKFTAFALSKNAKVLVHCIPAGISRSVTLVAATLKKNMNTTEALSHINER 164  
 OY 105 RSCANPNVGFQROLOEFER-----HEVHQYRQWLKEEXGE 139  
 DB 165 RSGISPNANFLROLRVYFECNYQLDRSLRPYRQWLFRRYGSD 205

RESULT 3  
 T32494  
 Hypothetical protein C05B10.1 - Caenorhabditis elegans  
 C: Species: Caenorhabditis elegans  
 C: Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C: Accession: T32494  
 R: Geisels, C.; Wamsley, P.  
 Submitted to the EMBL Data Library, December 1997  
 A: Description: The sequence of C. elegans cosmid C05B10.  
 A: Reference number: 221178  
 A: Accession: T32494  
 A: Status: preliminary; translated from GB/EMBL/DBDJB  
 A: Molecule type: DNA  
 A: Residues: 1-365 <GET>  
 A: Cross-references: EMBL:AF036665; PIDN: AAB88308.1; GSPDB: GNO0022; CESP: C05B10.1  
 A: Experimental source: strain Bristol N2; clone C05B10  
 C: Genetics:  
 A: Gene: CESP: C05B10.1  
 A: Map position: 4  
 A: Introns: 22/2; 46/2; 74/3; 163/2; 204/2; 248/1; 287/2

Query Match 21.6%; Score 194; DB 2; Length 365;  
 Best Local Similarity 31.4%; Pred. No. 8.4e-12;  
 Matches 43; Conservative 27; Mismatches 47; Indels 20; Gaps 2;

OY 7 KILPGLYIGNFKDARDAEQLSKNKYTHILSV-----HDSPLGLCW-----R 46  
 DB 183 KLTNELYIGNMARTAKNRDVLKYSISHVINVTSLPNTFEEDPNRRYLRLISADNASHL 242  
 OY 47 TRHFKESIKFTHECRLGESCTVHCLAGVSRVTLVIATVTVDFGMDALHTVRAGS 106  
 DB 243 TKFEFSAISFIDDARRNDACVHCLAGISRSVTLICLAYLTKTENCITLDSAEVWQKRNA 302  
 OY 107 CANPNVGFQROLOEFER 123  
 DB 303 STAPNFHEMGOLTDYER 319

RESULT 4  
 I38890  
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human  
 N: Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hVH-3;  
 C: Species: Homo sapiens (man)  
 C: Date: 16-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 11-Jun-1999  
 C: Accession: I38890; A55313  
 R: Kwak, S.P.; Dixon, J.E.  
 J. Biol. Chem. 270, 1156-1160, 1995  
 A: Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regulated  
 A: Reference number: A55432; MUID: 95138103

A: Accession: I38890  
 A: Molecule type: mRNA  
 A: Residues: 1-384 <RES>  
 A: Cross-references: EMBL: U16996; NID: g642012; PIDN: AAB06261.1; PID: g642013  
 A: Experimental source: placenta  
 R: Ishibashi, T.; Bottaro, D.P.; Michieli, P.; Kelley, C.A.; Aaronson, S.A.  
 J. Biol. Chem. 269, 29897-29902, 1994  
 A: Title: A novel dual specificity phosphatase induced by serum stimulation and heat  
 A: Reference number: A55313; MUID: 95050849  
 A: Accession: A55313

A: Molecule type: mRNA  
 A: Residues: 1-8, 'GHV', 12-70, 'R', 72-104, 'F', 107-362, 'RCLPTQSSSAALMORPAPAKTGEESAQP'  
 A: Cross-references: GB: U15932; NID: g606971; PIDN: AAA64693.1; PID: g606972  
 A: Experimental source: mammary epithelial cells  
 C: Genetics:  
 A: Gene: GDB: DUSP5  
 A: Cross-references: GDB: 385447  
 A: Map position: 10q25-10q25  
 C: Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specific  
 C: Keywords: phosphoprotein; phosphoric monoester hydrolase  
 F: 186-317/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1  
 F: 263/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F: 263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 20.7%; Score 186; DB 1; Length 384;  
 Best Local Similarity 33.6%; Pred. No. 5.7e-11;  
 Matches 46; Conservative 21; Mismatches 52; Indels 18; Gaps 1;

OY 4 GNNKILPGLYIGNFKDARDAEQLSKNKYTHILSVHDSFGLCWRT----- 48  
 DB 178 GVEIILPFLYIGSAVHAKCEFLANLHITALLNVSRRTSACMTHLHYKWIPEVDSHTAD 237  
 OY 49 ---HFKESIKFTHECRLGESCTVHCLAGVSRVTLVIATVTVDFGMDALHTVRAGR 105  
 DB 238 ISSHFQEAIDFIDCVREKGGKVLVHCAGISRSPTICMAYLTKTKQFLKFAFYIKORR 297  
 OY 106 SCANPNVGFQROLOEFE 122  
 DB 298 SWVSNFPGFMOQLLOYE 314

RESULT 5  
 T48906  
 protein-tyrosine-phosphatase (EC 3.1.3.48) [imported] - Chlamydomonas eugametos  
 C: Species: Chlamydomonas eugametos  
 C: Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 21-Jul-2000  
 C: Accession: T48906  
 R: Harting, M.A.; Siderius, M.; Joank, C.; Hirt, H.; Walton, K.M.; Musgrave, A.  
 Plant J. 7, 981-988, 1995  
 A: Title: Tyrosine phosphatase signalling in a lower plant: cell-cycle and oxidative  
 A: Reference number: 225005; MUID: 95323001  
 A: Accession: T48906  
 A: Status: preliminary; translated from GB/EMBL/DBDJB  
 A: Molecule type: mRNA  
 A: Residues: 1-276 <HAR>  
 A: Cross-references: EMBL: X77938; NID: g992593; PIDN: CAAS4910.1; PID: g992594  
 A: Experimental source: strain UTEX 10  
 C: Genetics:  
 A: Gene: VH-PTP13  
 C: Keywords: phosphoric monoester hydrolase

Query Match 20.6%; Score 185; DB 2; Length 276;  
 Best Local Similarity 33.3%; Pred. No. 4.8e-11;  
 Matches 52; Conservative 22; Mismatches 52; Indels 30; Gaps 5;  
 OY 8 ILPG-LYIGNFKDARDAEQLSKNKYTHILSVHDS-----PGLCWRTAHFKESI----- 54  
 DB 90 IVPKLLISSCEVESSLLKLGVTHTLQGEELKPSHPG---RTYLLSLPTIDMEGOD 146  
 OY 55 -----KTHFCRLRGESCTVHCLAGVSRVTLVIATVTVDFGMDALHTVRAGR 105

Db 147 IVALPSCFQLOQAGSGVCLVHCLAGISRSASVVIAYLMTQGMPTEARAVRRAR 206  
QY 106 SCANPNVGFQROLOEFKEHVEHYQRYWLKKEEYGSPP 141  
Db 207 SKVYPNTGFTLQLOELDLRLRESGAIQW-----GDP 237

## RESULT 6

149365  
protein tyrosine phosphatase - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I49365  
R:Wishart, M.J.; Denu, J.M.; Williams, J.A.; Dixon, J.E.  
J. Biol. Chem. 270, 26782-26785, 1995  
A:Title: A single mutation converts a novel phosphotyrosine binding domain into a dual-S  
A:Reference number: I49364; MUID:96070766  
A:Accession: I49365  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-223 <RES>  
A:Cross-references: EMBL:U034973; NID:g1063624; PIDN:AAA87037.1; PID:g1063626  
C:Genetics:  
A:Introns: 168/3  
C:Superfamily: VHL-type dual specificity phosphoprotein phosphatase homology  
F:36-174/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

Query Match 20.5%; Score 183.5; DB 2; Length 223;  
Best Local Similarity 33.8%; Pred. No. 5.3e-11;  
Matches 50; Conservative 18; Mismatches 53; Indels 27; Gaps 3;

QY 1 MGNGMKNKILPGLYIGNFEDARDAEO--LSKNKVTHT----- 34  
Db 25 MRREMQEVLPLGLPGYSSAMKSKLPILQKGITHIICIRONIEANFIKPNFQQLFRYL 84  
QY 35 LSVHDSPLGCRTRHFKFSIKFIHCCRLGSGCLVHCLAGVSRVTLVYIMTVTDGFW 94  
Db 85 LDIDNP-VENIIRFPMTKEFDISLQNGKVLVHGNAGISRSAAFYIAYIMETFGMKY 143  
QY 95 EDALHTVAGRSCANPNVGFQROLOEFE 122  
Db 144 RDAFYVQERRFCINPNAGFYHLOEYE 171

## RESULT 7

T03074  
dual specificity phosphoprotein phosphatase homolog - Chilo iridescent virus  
C:Species: Chilo iridescent virus  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 08-Oct-1999  
C:Accession: T03074  
R:Bahr, U.; Tioda, C.A.; Darai, G.  
Virus Genes 15, 235-245, 1997  
A:Title: The DNA sequence of Chilo iridescent virus between the genome coordinates 0.101  
A:Reference number: Z14834; MUID:98114193  
A:Accession: T03074  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-142 <BAH>  
A:Cross-references: EMBL:AF003534; NID:g2738385; PIDN:AAB94448.1; PID:g2738421

Query Match 20.3%; Score 182; DB 2; Length 142;  
Best Local Similarity 31.7%; Pred. No. 4.4e-11;  
Matches 44; Conservative 23; Mismatches 52; Indels 20; Gaps 3;

QY 7 KILPLTYGNFRKARDADQLSKNKVTHLSVHDSPLG--CQRT-----RHFKEKI 54  
Db 5 KIVENLYIGNIQNGIRHSNYGFKITINLRNNQIGIPYWINIDSSSDLYSHLOKVT 64  
QY 55 KFIHCRLRGESCLVHCLAGVSRVTLVYIMTVTDGFWEDALHTVAGRSCANPNVGF 114  
Db 65 TLIHDSINGKNVLYVHCAGISRSATVYIAYIMRSKRYSLDARFVKKKRSITFPNAGF 124

QY 115 QROLOEFKEHVEHYQRYWL 133  
Db 125 IKOLAQFER-----WL 135

## RESULT 8

149364  
protein tyrosine phosphatase - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: I49364  
R:Wishart, M.J.; Denu, J.M.; Williams, J.A.; Dixon, J.E.  
J. Biol. Chem. 270, 26782-26785, 1995  
A:Title: A single mutation converts a novel phosphotyrosine binding domain into a du  
A:Reference number: I49364; MUID:96070766  
A:Accession: I49364  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-205 <RES>  
A:Cross-references: EMBL:U034973; NID:g1063624; PIDN:AAA87036.1; PID:g1063625  
C:Superfamily: VHL-type dual specificity phosphoprotein phosphatase homology  
F:36-174/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

Query Match 19.3%; Score 173.5; DB 2; Length 205;  
Best Local Similarity 30.3%; Pred. No. 4.9e-10;  
Matches 53; Conservative 20; Mismatches 61; Indels 41; Gaps 5;

QY 1 MGNGMKNKILPGLYIGNFEDARDAEO--LSKNKVTHT----- 34  
Db 25 MRREMQEVLPLGLPGYSSAMKSKLPILQKGITHIICIRONIEANFIKPNFQQLFRYL 84  
QY 35 LSVHDSPLGCRTRHFKFSIKFIHCCRLGSGCLVHCLAGVSRVTLVYIMTVTDGFW 94  
Db 85 LDIDNP-VENIIRFPMTKEFDISLQNGKVLVHGNAGISRSAAFYIAYIMETFGMKY 143  
QY 95 EDALHTVAGRSCANPNVGFQROLOEFKEHVEHYQRYWLKKEEYGSPP--PLQDAE 146  
Db 144 RDAFYVQERRFCINPNAGF-----VHQLQLWLSNNSARSAPLPKQR 187

## RESULT 9

S31304  
protein-tyrosine-phosphatase (EC 3.1.3.48) YVH1 - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YIR026c  
C:Species: Saccharomyces cerevisiae  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 12-Nov-1999  
C:Accession: S31304; S48488; S28650  
R:Guan, K.; Hakes, D.J.; Wang, Y.; Park, H.D.; Cooper, T.G.; Dixon, J.E.  
Proc. Natl. Acad. Sci. U.S.A. 89, 12175-12179, 1992  
A:Title: A yeast protein phosphatase related to the vaccinia virus VHL phosphatase 1  
A:Reference number: S31304; MUID:93101690  
A:Accession: S31304  
A:Molecule type: DNA  
A:Residues: 1-364 <GUA>  
A:Cross-references: EMBL:L04673; NID:g172167; PIDN:AAA34874.1; PID:g172168  
R:Rowley, K.  
submitted to the EMBL Data Library, October 1994  
A:Reference number: S48478  
A:Accession: S48488  
A:Molecule type: DNA  
A:Residues: 1-364 <ROW>  
A:Cross-references: GB:Z47047; EMBL:Z38061; NID:g603997; PID:g763371; GSPDB:GMO0009;  
R:Buckholz, R.G.; Cooper, T.G.  
Yeast 7, 913-923, 1991  
A:Title: The allantoinase (DAL1) gene of Saccharomyces cerevisiae.  
A:Reference number: S28649; MUID:92206070  
A:Accession: S28650  
A:Molecule type: DNA  
A:Residues: 1-67,69-197,'AI' <BUC>  
A:Cross-references: EMBL:M69294  
C:Genetics:

A:Gene: SGD:YWH1; MIPS:YIR026C  
 A:Cross-references: SGD:S0001465; MIPS:YIR026C  
 A:Map position: 9R  
 C:Superfamily: Saccharomycetes protein-tyrosine-phosphatase YWH1, VHL-type dual specificity  
 C:Keywords: phosphoric monoester hydrolase; tyrosine-specific phosphatase  
 F:19-171/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL>

Query Match 19.0%; Score 170; DB 1; Length 364;  
 Best Local Similarity 24.9%; Pred. No. 2.2e-09;  
 Matches 51; Conservative 31; Mismatches 65; Indels 58; Gaps 6;

OY 5 MNKILPGLYIGNFKRDAEQL-SKNKVTHILSV-----HDSPG 42  
 DB 12 VTRILGIVLYGIRPIIDHRPLGAEPNTHILSVIKFOYIPEYLIRKGYTIKNPIIDDD 71  
 OY 43 LCKRTHFKESIKFIEHCRLRGSECLVHCLAGVSRSTLVIAV 85  
 DB 72 VYDVLIQYFDETNRFIDQCLFPNEVEYSPRLVDFKKRPGKGAFAHCOAGLSRSFTIVAY 131  
 OY 86 IMVTDFGWDALHTVRAGSCANPNVGFQROLOEFK-----HEVHOYRGM-LKEEY 137  
 DB 132 LMRVYGLSLSMAMHAYKRRKPSYEPNENFMDLHLFEKMGDFVDPNPAYKQWKLQSI 191  
 OY 138 GESP-----LDAAEAKNI 151  
 DB 192 KIDPSGSELVNSGMEKDESSQDL 216

RESULT 10  
 T21380  
 hypothetical protein F26A3.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T21380  
 R:McMurray, A.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: 219415  
 A:Accession: T21380  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-226 <MIL>  
 A:Cross-references: EMBL:278419; PIDN:CAB01700.1; GSPDB:GN00019; CESP:F26A3.4  
 A:Experimental source: clone F26A3  
 C:Genetics:  
 A:Gene: CESP:F26A3.4  
 A:Map position: 1  
 A:Introns: 117/2; 150/3; 186/3

Query Match 18.7%; Score 167.5; DB 2; Length 226;  
 Best Local Similarity 28.6%; Pred. No. 2.2e-09;  
 Matches 48; Conservative 33; Mismatches 58; Indels 29; Gaps 5;

OY 5 MNKILPGLYIGNFKRDAEQLSKNKVTHILSVHD-----SPGLCWTR----- 48  
 DB 14 MSEIYPGLEICGV-SALSDEKKKIRTHIINATTEVPLNSLGIQRIKMLEPTPOTY 72  
 OY 49 ---HFKESTIKFIEHCRLRGSECLVHCLAGVSRSTLVIAVIMVTDFGWDALHTVRAGR 105  
 DB 73 IYRPHLELQSDQIQALIDAGGKVLVHCVAGVSRASISICLAFILKYCRNLRVHLMKSKR 132  
 OY 106 SCANPNVGFQROLOEFKHEVHOYQWMLKEEYGESPL-QDAEAKNII 152  
 DB 133 SMVRNPLGFWRQLIAYEON-----VKENGVSLVRDEKQPOL 172

RESULT 11  
 T16056  
 hypothetical protein F13D11.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 23-Mar-2001  
 C:Accession: T16056

R:Fulton, L.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid F13D11.  
 A:Reference number: S69020  
 A:Accession: T16056  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-186 <FNU>

A:Cross-references: EMBL:U40939; NID:q1072175; PID:q1072176; PIDN:AAA81700.1; CESP:F  
 C:Genetics:  
 A:Gene: CESP:F13D11.3  
 A:Introns: 30/3; 57/3; 85/3; 125/3; 172/3

Query Match 18.3%; Score 164; DB 2; Length 186;  
 Best Local Similarity 39.0%; Pred. No. 3.9e-09;  
 Matches 30; Conservative 19; Mismatches 28; Indels 0; Gaps 0;

OY 47 TRHFESIKFIEHCRLRGSECLVHCLAGVSRSTLVIAVIMVTDFGWDALHTVRAGRS 106  
 DB 69 TQFFEVYVYIEDAKQGGHNVITYCAAGVSRSATILYIMVTENLSLEAYLYQNVYRP 128  
 OY 107 CANPNVGFQROLOEFK 123  
 DB 129 IISPNIGFWRMIDFEK 145

RESULT 12  
 S29090  
 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human  
 N:Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase,  
 C:Species: Homo sapiens (man)  
 C>Date: 25-Feb-1994 #sequence\_revision 02-May-1994 #text\_change 11-Jun-1999  
 C:Accession: S29090; A53052  
 R:Keyse, S.M.; Emslie, E.A.  
 Nature 359, 644-647, 1992  
 A:Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyro  
 A:Reference number: S29090; MUID:93024952  
 A:Accession: S29090  
 A:Molecule type: mRNA  
 A:Residues: 1-367 <KEY>  
 A:Cross-references: EMBL:X68277; NID:q29980; PIDN:CAA48338.1; PID:q29981  
 R:Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.  
 J. Biol. Chem. 269, 3596-3604, 1994  
 A:Title: Isolation and characterization of a human dual specificity protein-tyrosine  
 A:Reference number: A53052; MUID:94148864  
 A:Accession: A53052  
 A:Molecule type: DNA  
 A:Residues: 1-367 <KMA>  
 A:Experimental source: Leukocyte  
 A:Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143  
 C:Genetics:  
 A:Gene: GDB:DU5P1; PTPN10  
 A:Cross-references: GDB:136197; OMIM:600714  
 A:Map position: 5q34-5q34  
 C:Superfamily: dual specificity phosphoprotein phosphatase 1; VHL-type dual specific  
 C:Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induc  
 F:181-312/Domain: VHL-type dual specificity phosphoprotein phosphatase homology <VHL  
 F:256/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 17.9%; Score 161; DB 1; Length 367;  
 Best Local Similarity 29.8%; Pred. No. 1.8e-08;  
 Matches 42; Conservative 20; Mismatches 53; Indels 26; Gaps 2;

OY 4 GMNKILPGLYIGNFKRDAEQLSKNKVTHILSV-----HDSP 41  
 DB 173 GPVELIPFLYIGSAVHASRKMDLALITALLINVSANCPNHEGHYQKSPVEDNHRKD 232  
 OY 42 GLCKRTHFKESIKFIEHCRLRGSECLVHCLAGVSRSTLVIAVIMVTDFGWDALHTV 101  
 DB 233 ISSW-----FNEAIDPIDSIKNAGRVFVHCQGISRSATICTLAYIMRTNRYKIDAEFEV 288

Search completed: February 12, 2002, 16:13:32  
Job time: 191 sec





DB 135 GDSITLTPNYLGSQIDSDTMDALDISVYINLSMTCPSKVCIEDKFNRIEVDNS 194  
 QY 47 -----TRHEKESIKFIEHCRLGESCLVHCLAGVSRYTLVAVIMTVDFGMDALHTY 101  
 DB 195 YQKLSYFFPMAYEFLEKCRACKCLITRCLASISPTLAISYIMKMKGSDDAVRY 254  
 QY 102 RAGSCANPNVGFQROLQEFK-----KHEVQYRQWLK-----DEYGESPL 142  
 DB 255 KERRPSISPNFNFMGQLDEYENLVLIKDHVLDVYQASRPHRHMDYYGPSDL 304

## RESULT 2

DUS9\_HUMAN  
 ID DUS9\_HUMAN STANDARD: PRT: 384 AA.

AC 099956;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 9 (EC 3.1.3.48) (EC 3.1.3.16)  
 DE (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE 4) (MAP KINASE PHOSPHATASE 4) (MKP-4).  
 GN DUSP9 OR MKP4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;  
 RX MEDLINE=97184169; PubMed=9030581;

RA Muda M., Boschart U., Smith A., Antonsom B., Allieron C.,  
 Chabert C., Camps M., Martin I., Ashworth A., Arkinstall S.;

RT "Molecular cloning and functional characterization of a novel  
 mitogen-activated protein kinase phosphatase, MKP-4.";

RL J. Biol. Chem. 272:5141-5151(1997).

- FUNCTION: INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK  
 FAMILY.

- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 PROTEIN TYROSINE + PHOSPHATE.

- SUBCELLULAR LOCATION: CYTOPLASMIC.

- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: Y08302; CAA69610.1; -.

DR MIM: 300134; -.

DR InterPro: IPR000340; DS\_phosphatase.

DR InterPro: IPR001763; Rhodanese\_domain.

DR InterPro: IPR000387; TYR\_phosphatase.

DR Pfam: PF00782; DSPC: 1.

DR SMART: SM00195; DSPC: 1.

DR SMART: SM00450; RHOD: 1.

DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; FALSE\_NEG.

DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.

DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.

DR HydroLase.

FT DOMAIN 203 384 PROTEIN-TYROSINE PHOSPHATASE.

FT ACT\_SITE 290 290 BY SIMILARITY.

FT SEQUENCE 384 AA; 41867 MW; F8598CA95AB379B7 CRC64;

Query Match 23.1%; Score 207; DB 1; Length 384;

Best Local Similarity 34.9%; Pred. No. 3.7e-14;

Matches 53; Conservative 22; Mismatches 47; Indels 30; Gaps 3;

QY 7 KILPLGLYGNFKDARDADQLSKNKVTHILSV-----HDSPL 43

DB 206 QILPNLYLSARSDANLESLAKIGIRYILNVPNLPNFEKNDFHYKQIPISDHNSQL 265  
 QY 44 CWTRHFKESIKFIEHCRLGESCLVHCLAGVSRYTLVAVIMTVDFGMDALHTY 103  
 DB 266 ---SRFPEALIEFIDALSQNGCVLVHCLAGVSRYTVAVIMQTLHSLNDADYDLVR 322  
 QY 104 GRSCANPNVGFQROLQEFK-----HEVQYRQ 131  
 DB 323 KSNISPNFNFMGQLDFFERSLRLERHSQ 354

## RESULT 3

DUS8\_HUMAN  
 ID DUS8\_HUMAN STANDARD: PRT: 625 AA.

AC Q13202;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 8 (EC 3.1.3.48) (EC 3.1.3.16)  
 DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH-5).  
 GN DUSP8 OR VHS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=fetal brain;  
 RX MEDLINE=96009533; PubMed=7561881;

RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.,  
 "HVH-5: a protein tyrosine phosphatase abundant in brain that  
 inactivates mitogen-activated protein kinase.";

RL J. Neurochem. 65:1823-1833(1995).

- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN  
 PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY  
 SIMILARITY).

- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 PROTEIN TYROSINE + ORTHOPHOSPHATE.

- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.

- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U27193; AAA83151.1; -.

DR MIM: 602038; -.

DR InterPro: IPR000340; DS\_phosphatase.

DR InterPro: IPR002965; P-rich\_extensn.

DR InterPro: IPR001763; Rhodanese\_domain.

DR InterPro: IPR000387; TYR\_phosphatase.

DR Pfam: PF00782; DSPC: 1.

DR PRINTS: PRO1217; PRICHEXTENS.

DR SMART: SM00195; DSPC: 1.

DR SMART: SM00450; RHOD: 1.

DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.

DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.

DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.

DR HydroLase: Nuclear protein.

FT DOMAIN 28 43 CH2 A DOMAIN.

FT DOMAIN 117 132 CH2 B DOMAIN.

FT DOMAIN 162 430 PROTEIN-TYROSINE PHOSPHATASE.

FT DOMAIN 310 550 PRO-RICH.

FT ACT\_SITE 246 246 BY SIMILARITY.

FT SEQUENCE 625 AA; 65840 MW; DCBEA14487219666 CRC64;



[illegible]

DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 KW Hydrolase.  
 FT NON\_TER 1 1  
 ACT SITE 192 192 BY SIMILARITY.  
 SQ SEQUENCE 280 AA; 30668 MW; 45D6F4A92F2BDDF CRC64;

Query Match 22.0%; Score 197; DB 1; Length 280;  
 Best Local Similarity 35.0%; Pred. No. 2.8e-13;  
 Matches 49; Conservative 22; Mismatches 43; Indels 26; Gaps 2;

QY 7 KILPGYIGNFKDARAEOLSKNKVTHLSV-----HDSPLG 43  
 DB 108 QILPYLYLGCADSTNLDVKGKGYKILNVTNPNLPAFEGGEFTYKQIPISDHWSNL 167  
 QY 44 CWRTFRHFKESIKFIHECRLRGESCIVHCLAGVSRTLVIAIYMTYTDGWEALHTYRA 103  
 DB 168 ---SQFFPAISFIDERSKKGCVLHVCAGISRSVTVVAILMQRNLSLNDAYDFYKR 224  
 QY 104 GRSCANPNVGFQROJEFK 123  
 DB 225 KRSNISPNFNFMGQLDFFER 244

RESULT 6  
 ID DUS7\_HUMAN STANDARD; PRT; 322 AA.  
 AC 016829;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 7 (EC 3.1.3.48) (EC 3.1.3.16)  
 DE (DUAL SPECIFICITY PROTEIN PHOSPHATASE PYST2) (FRAGMENT).  
 GN DUSP7 OR PYST2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=96312959; PubMed=8670865;  
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;  
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyst1,  
 RL a novel cytosolic dual-specificity phosphatase.";  
 EMBO J. 15:3621-3632(1996).  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.ch](mailto:license@isb.ch)).  
 CC -----  
 CC EMBL: X93921; CA63814.1;  
 DR HSSP: P51452; 1YHR.  
 DR MIM: 602749;  
 DR InterPro: IPR000340; DS\_phosphatase.  
 DR InterPro: IPR000387; TYR\_phosphatase.  
 DR Pfam: PF00782; DSPC; 1.  
 DR SMART: SM00195; DSPC; 1.  
 DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; FALSE\_NEG.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 KW Hydrolase.  
 FT NON\_TER 1 1  
 ACT\_SITE 234 234 BY SIMILARITY.

SQ SEQUENCE 322 AA; 35449 MW; E89B1C2ABB2E75DE CRC64;

Query Match 22.0%; Score 197; DB 1; Length 322;  
 Best Local Similarity 35.0%; Pred. No. 3.3e-13;  
 Matches 49; Conservative 22; Mismatches 43; Indels 26; Gaps 2;

QY 7 KILPGYIGNFKDARAEOLSKNKVTHLSV-----HDSPLG 43  
 DB 150 QILPYLYLGCADSTNLDVKGKGYKILNVTNPNLPAFEGGEFTYKQIPISDHWSNL 209  
 QY 44 CWRTFRHFKESIKFIHECRLRGESCIVHCLAGVSRTLVIAIYMTYTDGWEALHTYRA 103  
 DB 210 ---SQFFPAISFIDERSKKGCVLHVCAGISRSVTVVAILMQRNLSLNDAYDFYKR 266  
 QY 104 GRSCANPNVGFQROJEFK 123  
 DB 267 KRSNISPNFNFMGQLDFFER 286

RESULT 7  
 ID DUS6\_HUMAN STANDARD; PRT; 381 AA.  
 AC 016828;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 6 (EC 3.1.3.48) (EC 3.1.3.16)  
 DE (MITOGEN-ACTIVATED PROTEIN KINASE PHOSPHATASE 3) (MAP KINASE  
 DE PHOSPHATASE 3) (MKP-3) (DUAL SPECIFICITY PROTEIN PHOSPHATASE PYST1).  
 GN DUSP6 OR PYST1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Foreskin;  
 RX MEDLINE=96312959; PubMed=8670865;  
 RA Groom L.A., Sneddon A.A., Alessi D.R., Dowd S., Keyse S.M.;  
 RT "Differential regulation of the MAP, SAP and RK/p38 kinases by Pyst1,  
 RL a novel cytosolic dual-specificity phosphatase.";  
 EMBO J. 15:3621-3632(1996).  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

RT	"A novel cytoplasmic dual specificity protein tyrosine phosphatase implicated in muscle and neuronal differentiation.";					
RL	Biol.	Chem.	271:3795-3802(1996).			
CC	-1-	FUNCTION:	INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK FAMILY. IMPLICATED IN MUSCLE AND NEURONAL DIFFERENTIATION.			
CC	-1-	CATALYTIC ACTIVITY:	PROTEIN TYROSINE PHOSPHATE + H <sub>2</sub> O = PROTEIN TYROSINE + PHOSPHATE.			
CC	-1-	SUBCELLULAR LOCATION:	CYTOPLASMIC.			
CC	-1-	SIMILARITY:	BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.			
CC	-----					
CC	This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/cc">http://www.isb-sib.ch/announce/cc</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).					
CC	-----					
DR	EMBL;	X94185;	CAAG3895.1; .			
DR	EMBL;	U42627;	AAB06202.1; .			
DR	HSSP;	P51452;	1VHR.			
DR	InterPro;	IPR000340;	DS_phosphatase.			
DR	InterPro;	IPR001763;	Rhodanese_domain.			
DR	InterPro;	IPR000387;	TYR_phosphatase.			
DR	Pfam;	PF00782;	DSPc; 1.			
DR	Pfam;	PF00581;	Rhodanese; 1.			
DR	SMART;	SMO0195;	DSPc; 1.			
DR	SMART;	SMO0450;	RHOD; 1.			
DR	PROSITE;	PS00383;	TYR_PHOSPHATASE_1; FALSE_NEG.			
DR	PROSITE;	PS50056;	TYR_PHOSPHATASE_2; 1.			
DR	PROSITE;	PS50054;	TYR_PHOSPHATASE_DUAL; 1.			
KW	Hydrolase.					
FT	DOMAIN	206	381			
FT	ACT_SITE	293	293			
SO	SEQUENCE	381 AA;	42318 MW; C31IED0CB68F2888 CMC64;			
	Query Match	21.9%;	Score 196;	DB 1;	Length 381;	
	Best Local Similarity	35.2%;	Pred. No. 5.2e-13;			
	Matches 50;	Conservative 24;	Mismatches 38;	Indels 30;	Gaps 4;	
QY	7	KILPELYIGNRKDAADAQLSKNKYTHLSV----	-----	HDSEGL	43	
Db	209	EILPELYGCAKDSINLVLEEFGIKYLLNTPNLPLEFNAGEFRYKQIPISDHMSNTL	268			
QY	44	CWRTRFKESIKETIECHRLRGESC--LVHCLAGVSRSYTVLAYITMTDPCGEDALHTY	101			
Db	269	---SQEPFEAFSEIDEA--RGANGCVLVHCLAGISRSYTVVAYLMOKLNLSMNDAYDI	323			
QY	102	RAGRSCAPNNVGFORLOEFEK	123			
Db	324	KMKKNISPNFMGQLDFER	345			
RESULT	9					
ID	DUS5_HUMAN	STANDARD:	PRT:	384 AA.		
AC	Q16690; Q12997;					
DT	01-NOV-1997 (Rel. 35, Created)					
DT	01-NOV-1997 (Rel. 35, Last sequence update)					
DT	20-AUG-2001 (Rel. 40, Last annotation update)					
DE	DUAL SPECIFICITY PROTEIN PHOSPHATASE 5 (EC 3.1.3.48) (EC 3.1.3.16)					
DE	(DUAL SPECIFICITY PROTEIN PHOSPHATASE HVH3).					
GN	DUSP5 OR VH3.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
OX	NCBI_TaxID=9606;					
NP	[1]					
RP	SEQUENCE FROM N.A.A.					
TISSUE	Tissue=Placenta;					
KX	MEDLINE=95138103; PubMed=7836374;					

RA Kwak S.P., Dixon J.E.;  
 RT "Multiple dual specificity protein tyrosine phosphatases are  
 RT expressed and regulated differentially in liver cell lines."  
 RL J. Biol. Chem. 270:1156-1160(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=95050849; PubMed=7961985;  
 RA Ishibashi T., Bottaro D.P., Michieli P., Kelley C.A.,  
 RA Aaronson S.A.;  
 RT "A novel dual specificity phosphatase induced by serum stimulation  
 RT and heat shock."  
 RL J. Biol. Chem. 269:29897-29902(1994).  
 RN [3]  
 RP REVISIONS.  
 RA Bottaro D.P.;  
 RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: REPLAYS PHOSPHATASE ACTIVITY TOWARD SEVERAL SUBSTRATES.  
 CC THE HIGHEST RELATIVE ACTIVITY IS TOWARD ERL1.  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U16996; AAB06261.1; -  
 DR EMBL: U15932; AAB64693.2; -  
 DR HSSP: P51452; 1YHR.  
 DR MIM: 603069; -  
 DR InterPro: IPR000340; DS-phosphatase.  
 DR InterPro: IPR001763; Rhodanese domain.  
 DR InterPro: IPR000387; TYR\_phosphatase.  
 DR Pfam: PF00782; DSPC; 1.  
 DR Pfam: PF00581; Rhodanese; 1.  
 DR SMART: SM00195; DSPC; 1.  
 DR SMART: SM00450; RHOD; 1.  
 DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
 DR PROSITE: PS00056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 KW Hydrolyase; Nuclear protein.  
 FT DOMAIN 24 39  
 FT DOMAIN 53 74  
 FT DOMAIN 79 82  
 FT DOMAIN 120 135  
 FT DOMAIN 180 384  
 FT ACI\_SITE 263 263  
 FT ACI\_SITE 263 263  
 FT CONFLICT 9 11  
 FT CONFLICT 71 71  
 FT CONFLICT 105 106  
 FT SEQUENCE 384 AA; 42107 MW; D2B726F7C0414306 CRC64;  
 SQ  
 Query Match 20.7%; Score 186; DB 1; Length 384;  
 Best Local Similarity 33.6%; Pred. No. 5.8e-12;  
 Matches 46; Conservative 21; Mismatches 52; Indels 18; Gaps 1;  
 Oy 4 GNMKILPGLYIGNFKDADAQLSKNKVTHTLSVHDSGLCWRR----- 48  
 Db 178 GPVELLPGLYIGSAVHASKCEFLNLTALTNVSRRTSEACMTHLHKWIPVEDSHAD 237  
 Oy 49 ---HFKESEIKFTHECRVLGECVAGSVSTVLTATYITDEFGMEDALHTVRGR 105  
 Db 238 ISSHFQEAIDFTDVKREKGVLVHCEAGISRPTICMAYIMKTKQFLKEAFDYIKOR 297  
 Oy 106 SCANPNVGFQROLOEFE 122

Db 298 SMVSPNFGFMGLQLOVE 314  
 I :||| || :||  
 RESULT 10  
 DUS5\_RAT  
 ID DUS5\_RAT STANDARD; PRT; 384 AA.  
 AC O54838;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 5 (EC 3.1.3.48) (EC 3.1.3.16)  
 DE (MAP-KINASE PHOSPHATASE CPG21).  
 GN DUSP5.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_Taxid=101116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar;  
 RX MEDLINE=96364306; PubMed=9699150;  
 RA Heyvont D., Ralther A., Bundman M., Lederfein D., Gabarrah A.,  
 RA Mangelus M., Silverman M.A., Kedar H., Naor C., Kornuc M., Hancock T.,  
 RA Seger R., Theill L.E., Nedivi E., Richter-Levin G., Citri Y.;  
 RT "Hippocampal plasticity involves extensive gene induction and multiple  
 RT cellular mechanisms."  
 RL J. Mol. Neurosci. 10:75-98(1998).  
 CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN  
 CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
 CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS THE CDC25 HOMOMOLOGY DOMAINS 2 A AND B (CH2  
 CC DOMAINS A AND B).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF013144; AAB94858.1; -  
 DR InterPro: IPR000340; DS-phosphatase.  
 DR InterPro: IPR001763; Rhodanese domain.  
 DR InterPro: IPR000387; TYR\_phosphatase.  
 DR Pfam: PF00782; DSPC; 1.  
 DR Pfam: PF00581; Rhodanese; 1.  
 DR SMART: SM00195; DSPC; 1.  
 DR SMART: SM00450; RHOD; 1.  
 DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
 DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 KW Hydrolyase; Nuclear protein.  
 FT DOMAIN 24 39  
 FT DOMAIN 53 74  
 FT DOMAIN 79 82  
 FT DOMAIN 120 135  
 FT DOMAIN 180 384  
 FT ACI\_SITE 263 263  
 FT ACI\_SITE 263 263  
 FT SEQUENCE 384 AA; 42094 MW; 56440698BD348700 CRC64;  
 SQ  
 Query Match 20.7%; Score 186; DB 1; Length 384;  
 Best Local Similarity 34.3%; Pred. No. 5.8e-12;  
 Matches 47; Conservative 19; Mismatches 53; Indels 18; Gaps 1;  
 Oy 4 GNMKILPGLYIGNFKDADAQLSKNKVTHTLSVHDSGLCWRR----- 48

DB 178 GPEVLIPFLVLSGASVSKCEFLANLHITALLNVRRTSACTHTLHKYMLPVEDSHFAD 237  
QY 49 ---HFKESTFTHIECHRLRGSCVHCLAGVSRSTVIAIVTYVDFGMDALHTVAGR 105  
DB 238 ISSHFOEALIDFCVREEGKVLVHCAGVSRSPITICMAVLMKORFLKEAFEYIKORR 297  
QY 106 SCANPNVGFORLOEPE 122  
DB 298 SVVSPNFGFMGOLLOYE 314  
RESULT 11  
PTP3\_CHLEU  
ID PTP3\_CHLEU STANDARD; PRT; 276 AA.  
AC 039491;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE 15-JUL-1998 (Rel. 36, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE (EC 3.1.3.48) (EC 3.1.3.16).  
GN YH-PP13.  
OS Chlamydomonas eugametos.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3053;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=UTEX 10;  
RC MEDLINE=95323001; PubMed=7599654;  
RA Haring M.A., Siderius M., Jonak C., Hirt H., Walton K.M.,  
RA Mosgrave A.;  
RT Tyrosine phosphatase signalling in a lower plant: cell-cycle and  
RT oxidative stress-regulated expression of the Chlamydomonas eugametos  
RT YH-PP13 gene.";  
RL Plant J. 7:981-988(1995).  
CC -1- FUNCTION: COULD BE INVOLVED IN TYROSINE PHOSPHATASE SIGNALLING  
CC PATHWAYS, HAVING MAP-KINASES AS SUBSTRATES.  
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -1- DEVELOPMENTAL STAGE: NONDIVIDING GAMETES DID NOT EXPRESS THE YH-  
CC PTP13 GENE WHEREAS SYNCHRONOUSLY DIVIDING VEGETATIVE CELLS ONLY  
CC EXPRESSED YH-PP13 IN THE EARLY G1-PHASE OF THE CYCLE.  
CC -1- INDUCTION: BY OXIDATIVE STRESS.  
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X77938; CAA54910.1; -  
DR HSSP: P51452; 1VHR.  
DR InterPro: IPR000340; DS\_phosphatase.  
DR InterPro: IPR000387; TYR\_phosphatase.  
DR Pfam: PF00782; DSPC; 1.  
DR SMART: SM00195; DSPC; 1.  
DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Hydrolyase.  
FT DOMAIN 89 276 CATALYTIC.  
FT ACT\_SITE 172 172  
FT MUTAGEN 172 172 C-S: INACTIVE.  
SQ SEQUENCE 276 AA; 30310 MM; 034EF6351E03381 CRC64;  
Query Match 20.6%; Score 185; DB 1; Length 276;  
Best Local Similarity 33.3%; Pred. No. 5e-12;  
Matches 52; Conservative 22; Mismatches 52; Indels 30; Gaps 5;

QY 8 ILPG-LYIGNFKDARDAPOLSKNKVTHLSVHDS-----PGLCWRTFRFKESI----- 54  
DB 90 IVPGLILSSCEVESSSLTKLGVTHLYGEELEKPSHPG---RFTYLSLPILDMGOD 146  
QY 55 -----KFTHIECHRLRGSCVHCLAGVSRSTVIAIVTYVDFGMDALHTVAGR 105  
DB 147 IVALLPSCGFLOQAQASGVCLVHCLAGISRSASVATIAMVMQGMYTEARAVRRAR 206  
QY 106 SCANPNVGFORLOEPEFKEHVEYQYRWLKEEYGESE 141  
DB 207 SKVYPNMGFTQLQLOELDRRESGAIQW-----GDTPE 237  
RESULT 12  
PVH1\_YEAST  
ID PVH1\_YEAST STANDARD; PRT; 364 AA.  
AC 002256;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE PROTEIN-TYROSINE PHOSPHATASE YVH1 (EC 3.1.3.48) (PTPASE YVH1).  
GN YVH1 OR YIR026C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=93101690; PubMed=1334559;  
RA Guan K., Hakes D.J., Wang Y., Park H.-D., Cooper T.G., Dixon J.E.;  
RT "A yeast protein phosphatase related to the vaccinia virus Vh1  
RT phosphatase is induced by nitrogen starvation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:12175-12179(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
RA Churcher C.M., Connor R., Cosey T., Dear S., Devlin K., Fraser A.,  
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
RA Louis E., Lye G., Moulé S., Moulé T., Odell C., Pearson D.,  
RA Rajandream M.A., Riles L., Rowley N., Skellon J., Smith V.,  
RA Walsh S.V., Whitehead S.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1-197 FROM N.A.  
RC MEDLINE=92206070; PubMed=1803816;  
RA Buckholz R.G., Cooper T.G.;  
RT "The allantoinase (DAL1) gene of Saccharomyces cerevisiae.";  
RL Yeast 7:913-923(1991).  
RN [4]  
RP ERRATUM.  
RA Buckholz R.G., Cooper T.G.;  
RL Yeast 8:239-239(1992).  
RN [5]  
RP SIMILARITY TO YH1.  
RC MEDLINE=93174465; PubMed=8438236;  
RA Guan K., Hakes D.J., Dixon J.E., Park H.-D., Cooper T.G.;  
RT "The yeast open reading frame encoding a dual specificity  
RT phosphatase.";  
RL Trends Biochem. Sci. 18:6-6(1993).  
CC -1- FUNCTION: MAY BE DIRECTLY INVOLVED IN SIGNAL TRANSDUCTION AND/OR  
CC CELL CYCLE REGULATION. IT IS NECESSARY FOR MAINTAINING GROWTH RATE  
CC OR SPORE GERMINATION. COULD SHOW BOTH ACTIVITY TOWARD TYROSINE-  
CC PROTEIN PHOSPHATE AS WELL AS WITH SERINE-PROTEIN PHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.  
CC -1- INDUCTION: BY NITROGEN STARVATION.  
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L04673; AAA34874.1; -;  
DR EMBL; Z38061; CAAB6186.1; -;  
DR EMBL; M69294; -; NOT\_ANNOTATED\_CDS.  
DR PIR; S31304; S31304.  
DR PIR; S28650; S28650.  
DR PIR; S48488; S48488.  
DR SGD; S0001465; YVH1.  
DR InterPro; IPR000340; DS\_phosphatase.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR Pfam; PF00782; DSPC; 1.  
DR SMART; SM00195; DSPC; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; FALSE\_NEG.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
KW Hydrolyase.  
FT DOMAIN 1 172 PROTEIN-TYROSINE PHOSPHATASE.  
FT ACT\_SITE 117 117 BY SIMILARITY.  
FT CONFLICT 71 71 MISSING (IN REF. 3).  
SQ SEQUENCE 364 AA; 41185 MW; 9E75C7C14353B43 CRC64;

Query Match 19.0%; Score 170; DB 1; Length 364;  
Best Local Similarity 24.9%; Pred. No. 2.6e-10;  
Matches 51; Conservative 31; Mismatches 65; Indels 58; Gaps 6;

QY 5 MNKILPGLYIGNFKDARAEOLSKNKVTHLSV-----HDSFG 42  
DB 12 VTRILGGLYIGGIRIIDRLPLGAEFNTHLSVTKFQYIPRYLLRKGTLLKNIPIDDD 71  
QY 43 LCMRTFRKESIKFTHECRLGE-----SCLVHCLAGVSRVTLVIAY 85  
DB 72 VTDVLQYDFETNRFDICLFNEVEYSPRLVDFKKKPGQGAFAVHCQGLSRVTFIYAY 131  
QY 86 IMTVDFPQMEDALHTVRAGRSANPNVGFQROLOFEK-----HEVHQYQOM-LKEEY 137  
DB 132 LMYRGLSLSMAMHAKVKKRKSVEPNENMDLHLEKMGDGFVDNPVAYKQMKLKOI 191  
QY 138 GESP-----LDAAEAKNI 151  
DB 192 KLDPGSELVNSGMFKDSESSQDL 216

RESULT 13  
DUSC\_HUMAN STANDARD: PRT: 340 AA.  
AC 09UN16;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 12 (EC 3.1.3.48) (EC 3.1.3.16)  
DE (DUAL SPECIFICITY TYROSINE PHOSPHATASE YVH1).  
GN DUSP12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RX MEDLINE=99377030; PubMed=10446167;  
RA Mada M., Manning E.R., Orth K., Dixon J.E.;  
RT "Identification of the human YVH1 protein-tyrosine phosphatase orthologue reveals a novel zinc binding domain essential for in vivo function.";  
RT J. Biol. Chem. 274:23991-23995(1999).  
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =  
CC PROTEIN TYROSINE + PHOSPHATE.  
CC -1- COFACTOR: BINDS 2 MOLES OF ZINC PER MOLE OF PROTEIN.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- TISSUE SPECIFICITY: UBIQUITOUS. HIGHEST EXPRESSION IN SPLEEN, TESTIS, OVARY, AND PERIPHERAL BLOOD LEUKOCYTES AND LOWER EXPRESSION IN LIVER AND LUNG.  
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF119226; AAD51134.1; -;  
DR MIM; 604835; -;  
DR InterPro; IPR000340; DS\_phosphatase.  
DR InterPro; IPR000387; TYR\_phosphatase.  
DR InterPro; IPR000822; Znf-C2H2.  
DR Pfam; PF00782; DSPC; 1.  
DR SMART; SM00195; DSPC; 1.  
DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; FALSE\_NEG.  
DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; UNKNOWN\_1.  
KW Hydrolyase; Zinc; Metal-binding.  
FT ACT\_SITE 115 115 BY SIMILARITY.  
SQ SEQUENCE 340 AA; 37687 MW; 56B52192B42C73EB CRC64;

Query Match 18.7%; Score 168; DB 1; Length 340;  
Best Local Similarity 29.3%; Pred. No. 3.8e-10;  
Matches 49; Conservative 27; Mismatches 59; Indels 32; Gaps 6;

QY 5 MNKILPGLYIGNFKDARAEOLSKNKVTHLSVHD-----SPGL-CWR----- 46  
DB 27 MEVQPGLYFGCAAAVAEPDHLREAGITAVLVDSPEPSFKAGPVEDMLRFVADLKP 86  
QY 47 ----TRHRESIKFTHECRLGESCIVHCLAGVSRVTLVIAYIMTVDFPQMEDALHTVR 102  
DB 87 ETDLSLHLDRCVAFIFGQAAEAGRAVLVHCHAGVSRVTAITFAIMKTQDLPEKAYEKIO 146  
QY 103 AGRSANPNVGFQROLOFEK--HEV-----HQYR-OMLKEEY 139  
DB 147 ILKPKAKNKEGFEMOLKLIQAMGYEVDISSAIYKRIQOKTEKYPE 193

RESULT 14  
DUSL\_HUMAN STANDARD: PRT: 367 AA.  
AC P28562;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DUAL SPECIFICITY PROTEIN PHOSPHATASE 1 (EC 3.1.3.48) (EC 3.1.3.16)  
DE (MAP KINASE PHOSPHATASE-1) (MKP-1) (PROTEIN-TYROSINE PHOSPHATASE CL100) (DUAL SPECIFICITY PROTEIN PHOSPHATASE VH1).  
GN DUSP1 OR PTPN10 OR MKP1 OR CL100 OR VH1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RX MEDLINE=93024952; PubMed=1406996;  
RA Keyes S.M., Emslie E.A.;  
RT "Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine phosphatase.";  
RT Nature 359:644-647(1992).  
CC -1- FUNCTION: DUAL SPECIFICITY PHOSPHATASE THAT DEPHOSPHORYLATES MAP KINASE ERK2 ON BOTH THR-183 AND TYR-185.



---





QY 104 GRSCANPVNGFOROLOEFKEHEVHOYROWLKEEGESPLDPAEAKNLLAAGILKFWAF 163  
DB 121 GRSCANPVNGFOROLOEFKEHEVHOYROWLKEEGESPLDPAEAKNLLAAGILKFWAF 180  
QY 164 LRRL 167  
DB 181 LRRL 184

RESULT 2  
Q99N11 PRELIMINARY; PRT; 184 AA.

AC 099N11:  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE DUAL SPECIFICITY PHOSPHATASE TS-DSP2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Aoyama K., Matsuda T., Aoki N.;  
RT "Molecular cloning of a novel dual specificity phosphatase TS-DSP2.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF237619; AAK15038.1;--  
SQ SEQUENCE 184 AA; 20997 MW; 64953325E8AB577 CRC64;

Query Match 88.2%; Score 791.5; DB 11; Length 184;  
Best Local Similarity 82.6%; Pred. No. 7.1e-76;  
Matches 152; Conservative 11; Mismatches 4; Indels 17; Gaps 2;

QY 1 MGNGMKNKILPGLYIGNFKDARDAEQLSKNKYTHILSVHDSFG-----LCW----- 45  
DB 1 MGSGMSQILPGLYIGNFKDARDAEQLSKNKYTHILSVHDTARPMLEGVKYLCTIPADRP 60  
QY 46 --RRHFKESTIKFHECHLRGESCCLVHCLAGVSRSVTLVIATIMVTDFGWDALHTYRA 103  
DB 61 QNLRHFKESTIKFHECHLRGESCCLVHCLAGVSRSVTLVIATIMVTDFGWDALHTYRA 120  
QY 104 GRSCANPVNGFOROLOEFKEHEVHOYROWLKEEGESPLDPAEAKNLLAAGILKFWAF 163  
DB 121 GRSCANPVNGFOROLOEFKEHEVHOYROWLKEEGESPLDPAEAKNLLAAGILKFWAF 180  
QY 164 LRRL 167  
DB 181 LRRL 184

RESULT 3  
Q9H1R2 PRELIMINARY; PRT; 275 AA.

AC 09H1R2:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE BA243J16.6 (NOVEL PROTEIN) (FRAGMENT).  
GN BA243J16.6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Brown A.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL1601175; CAC10008.1;--  
DR InterPro; IPR000340; DS\_phosphatase.  
DR Pfam; PF00782; DSPC; 1.  
DR SMART; SM00195; DSPC; 1.

DR PROSITE; PS0054; TYR PHOSPHATASE\_DUAL; 1.  
FT NON TER 1  
SQ SEQUENCE 275 AA; 30129 MW; E1F480D6BD3D3F59 CRC64;

Query Match 45.1%; Score 404.5; DB 4; Length 275;  
Best Local Similarity 43.5%; Pred. No. 9.7e-35;  
Matches 80; Conservative 26; Mismatches 39; Indels 39; Gaps 3;

QY 1 MGNGMKNKILPGLYIGNF-----KDARDAEQLSKNKYTHILSVHDSFGLCW 45  
DB 33 MGNGMKNKILPGLYIGNFHPASQIGSSILFLSDAKDLDQGRKNTHTSHIESPQPLL 92  
QY 46 R-----TRHFKESTIKFHECHLRGESCCLVHCLAGVSRSVTLVIATIMT 88  
DB 93 QDITTLRLPVADTPPEVPIKHKKECINFHCRLNGKCLVLC-----TTIVATYVMT 145

QY 89 VTDFGWDALHTVRAGSCANPVNGFOROLOEFKEHEVHOYROWLKEEGESPLDPAEERA 148  
DB 146 VTGLGMRDYLVAIKATRPANPNPGRQOLEFGNASSQKLRLROLEERGESPFDRDEEL 205

QY 149 KNIL 152  
DB 206 RALL 209

RESULT 4  
Q9VU80 PRELIMINARY; PRT; 608 AA.

AC 09VU80:  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE CG10089 PROTEIN.  
GN CG10089.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
SX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Centler A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Dudin K.J., Evangelista C.C., Ferrara C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Plessner J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

DR InterPro: IPR000387; TYR\_phos  
DR Pfam: PF00782; DSPC; 1.  
DR Pfam: PF00581; Rhodanese; 1.  
DR SMART: SM00195; DSPC; 1.



QY 105 RSCANPNVGFQROLOEFK 123  
 ID 443 RPIISPNLNMGLIEEF 461

RESULT 9  
 ID 09TY00 PRELIMINARY; PRT; 221 AA.

AC 09TY00; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE C16A3.2. PROTEIN.  
 GN C16A3.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;

RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).

RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Favell A.;  
 RT "The sequence of C. elegans cosmid C16A3.";  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U41534; AAD31049.1; .  
 DR HSSP; Q16828; IMKP.  
 DR InterPro; IPR000340; DS\_phosphatase.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF00782; DSPC; 1.

DR SMART; SM00195; DSPC; 1.  
 DR PROSITE; PS00383; TYR\_PHOSPHATASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 SQ SEQUENCE 221 AA; 25362 MW; 2ECF68CF1452FA7F CRC64;

Query Match 23.2%; Score 208.5; DB 5; Length 221;  
 Best Local Similarity 27.6%; Pred. No. 4e-14;  
 Matches 47; Conservative 35; Mismatches 63; Indels 25; Caps 2;

QY 5 MNKPLGLYIGNFKDARDAEOLSKNKVTHILSVHDSPLGCMRTFH----- 49  
 ID 44 ISELPLVLSGRVTSQNSSELEKKNITVYINSDREVNVYKNNQKFKKNRYAASDTA 103  
 DB 50 -----FKSIFFIHCEKRLGSCLYHCLAGVSRSTVLYAIIMVTYTDGMEALHTVR 102  
 DB 104 SAKFDGIEAVRIIHDSRKEGVLVHCLGVSRSATLVAFYLSALSIWMDAVDFIH 163  
 QY 103 AGRSCANPNVGFQROLOEFKHEVHYROWLKEEYGESPLDDEAKNIL 152  
 ID 164 HRFPSANPNVGFHLQLVYSTTKAKFRNOL---ISERCLMRSDXDTI 210

RESULT 10  
 ID 09M8K7 PRELIMINARY; PRT; 167 AA.  
 AC 09M8K7; 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE PUTATIVE DUAL-SPECIFICITY PROTEIN PHOSPHATASE.  
 GN F28L1.5.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,  
 RA Rongling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RT "Arabidopsis thaliana chromosome III BAC F28L1 genomic sequence.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC018907; AAF30304.1; .  
 DR InterPro; IPR000340; DS\_phosphatase.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF00782; DSPC; 1.  
 DR SMART; SM00195; DSPC; 1.  
 DR PROSITE; PS50056; TYR\_PHOSPHATASE\_2; 1.  
 DR PROSITE; PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 SQ SEQUENCE 167 AA; 18431 MW; 57D722910B79A900 CRC64;

Query Match 23.1%; Score 207; DB 10; Length 167;  
 Best Local Similarity 34.3%; Pred. No. 4.1e-14;  
 Matches 49; Conservative 22; Mismatches 48; Indels 24; Caps 2;

QY 2 GNGMKILPLGLYIGNFKDARDAEOLSKNKVTHILSV-----HDS 40  
 ID 22 GKDLSEIIOGFLIGSVAAENKKNKDFIKSSNITHLVAVAPPPDDPVYKVEVDRSE 81  
 DB 41 PGLCMRTFHFKESIKFIHCEKRLGSCLYHCLAGVSRSTVLYAIIMVTYTDGMEALHT 100  
 DB 82 TDL---TYFDECYSEFIDQAIQSGGVLVHCFMGSRSSTVLYAVLMKKHGMGFSKAMEL 138  
 QY 101 VRAGSCANPNVGFQROLOEFK 123  
 ID 139 VRSRRHQAVPNPVGFIQLOEFK 161

RESULT 11  
 ID 09ZR37 PRELIMINARY; PRT; 198 AA.  
 AC 09ZR37; 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE DSPTP1 PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99155801; PubMed=10036776;  
 RA Gupta R., Huang Y., Kleber J., Luan S.;  
 RT "Identification of a dual-specificity protein phosphatase that  
 inactivates a MAP kinase from Arabidopsis.";  
 RL Plant J. 16:581-589(1998).

RP SEQUENCE FROM N.A.  
 RA Gupta R.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y18620; CAA77232.1; .  
 DR HSSP; Q16828; IMKP.  
 DR InterPro; IPR000340; DS\_phosphatase.  
 DR InterPro; IPR000387; TYR\_phosphatase.  
 DR Pfam; PF00782; DSPC; 1.  
 DR SMART; SM00195; DSPC; 1.

DR PROSITE: PS50056; TYR\_PHOSPHATASE.2; 1.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 SO SEQUENCE 198 AA; 22017 MW; EBFIC98A17E6450 CRC64;

Query Match 22.9%; Score 205; DB 10; Length 198;  
 Best Local Similarity 37.8%; Pred. No. 8.2e-14;  
 Matches 51; Conservative 14; Mismatches 44; Indels 26; Gaps 2;

QY 11 GLYIGNEFDARDADQLSKNKVTHILSVHDSFGLCWRTH----- 49  
 DB 57 GLYIGSVAAASNKVNLKSYNTHILTVASS-----LRPAHPDFYKVVYVDKEDTNLEM 112  
 QY 50 -FKESIKFIHECRLEGESCLVHCLAGVRSVTLVATYMTYTDGEMDALHTVRAGRSCA 108  
 DB 113 YFDECVDFIDAKRGGGSLVHCFVGKRSRVTIVAYILMKRHGMLTALQALHVKSKRPVA 172  
 QY 109 NPNVGFOROLOFEK 123  
 DB 173 SPNAGFIROLQDLEK 187

RESULT 12  
 09LUG6 PRELIMINARY; PRT: 198 AA.

AC 09LUG6; 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE DUAL-SPECIFICITY PROTEIN PHOSPHATASE-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA.  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA.  
 RX MEDLINE=20277480; PubMed=10819329;  
 RA Nakamura Y.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 features of the regions of 4,504,864 bp covered by sixty P1 and TAC  
 RT clones";  
 RT DNA Res. 7:131-135(2000).  
 RL EMBL: AB023036; BAB02780.1; -  
 DR InterPro: IPR000340; DS\_phosphatase.  
 DR InterPro: IPR000387; TYR\_phosphatase.  
 DR Pfam: PF00782; DSPc; 1.  
 DR SMART: SM00195; DSPc; 1.  
 DR PROSITE: PS50056; TYR\_PHOSPHATASE.2; 1.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 SO SEQUENCE 198 AA; 22113 MW; 815BCBE3BF696A CRC64;

Query Match 22.9%; Score 205; DB 10; Length 198;  
 Best Local Similarity 37.8%; Pred. No. 8.2e-14;  
 Matches 51; Conservative 14; Mismatches 44; Indels 26; Gaps 2;

QY 11 GLYIGNEFDARDADQLSKNKVTHILSVHDSFGLCWRTH----- 49  
 DB 57 GLYIGSVAAASNKVNLKSYNTHILTVASS-----LRPAHPDFYKVVYVDKEDTNLEM 112  
 QY 50 -FKESIKFIHECRLEGESCLVHCLAGVRSVTLVATYMTYTDGEMDALHTVRAGRSCA 108  
 DB 113 YFDECVDFIDAKRGGGSLVHCFVGKRSRVTIVAYILMKRHGMLTALQALHVKSKRPVA 172  
 QY 109 NPNVGFOROLOFEK 123

DB 173 SPNAGFIROLQDLEK 187

RESULT 13  
 013632 PRELIMINARY; PRT: 330 AA.

AC 013632; 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PROTEIN-TYROSINE PHOSPHATASE YVH1.  
 GN P1040 OR SPAC17A3.06.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972 H-;  
 RA Kushida N., Yamazaki S., Tanaka T., Jinno K., Haikawa Y., Yamazaki J.,  
 RA Yamamoto S., Sekine M., Oguchi A., Nagai Y., Sakai M., Aoki K.,  
 RA Ogura K., Otsuka R., Kudo Y., Yanagida M., Machida M., Zhang M.Q.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA Wood V., Skellon J., Churcher C.M., Rajandream M.A., Barrell B.G.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB004537; BAA21420.1; -  
 DR EMBL: AL109652; CAB51765.1; -  
 DR HSRP: Q16828; IMKP.  
 DR InterPro: IPR000340; DS\_phosphatase.  
 DR InterPro: IPR000387; TYR\_phosphatase.  
 DR Pfam: PF00782; DSPc; 1.  
 DR SMART: SM00195; DSPc; 1.  
 DR SMART: SM00192; PTPC; DSPc; 1.  
 DR PROSITE: PS50056; TYR\_PHOSPHATASE.2; 1.  
 DR PROSITE: PS50054; TYR\_PHOSPHATASE\_DUAL; 1.  
 SO SEQUENCE 330 AA; 38006 MW; DBF64E7FAEBA1E84 CRC64;

Query Match 22.5%; Score 202; DB 3; Length 330;  
 Best Local Similarity 32.3%; Pred. No. 3.1e-13;  
 Matches 52; Conservative 26; Mismatches 59; Indels 24; Gaps 4;

QY 3 NGMNKILPGLYIGNEFDARDADQLSKNKVTHILSVHD--SPGL-----CWR----- 46  
 DB 45 NOLSELKRLIYSSWKKTASLSTYDCKGIDYLSAMSINPNLSVPEQGHLMQIEDSSQ 104  
 QY 47 --TRFKESIKFIHECRLEGESCLVHCLAGVRSVTLVATYMTYTDGEMDALHTVRAG 104  
 DB 105 NILQYFEKSNKFIYALSKNAKLVHCFAGISRSVTLVAYILMKRNMTBETALSHINER 164  
 QY 105 RSCAMPNNGFOROLOFEK-----HEVHQYQWMLKEEYGE 139  
 DB 165 RSGISPMNMFRLQRLVFECCYNQDLRSLRPYQWMLFRRYGD 205

RESULT 14  
 091663 PRELIMINARY; PRT: 378 AA.

AC 091663; 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE DUAL SPECIFICITY PROTEIN PHOSPHATASE ? (EC 3.1.3.48) (EC 3.1.3.16)  
 OS (MAP KINASE PHOSPHATASE X17C).  
 DE Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA MEDLINE=95319535; PubMed=7541116;
RT Umhauer M., Marshall C.J., Mason C.S., Old R.W., Smith J.C.;
RL "Mesoderm induction in Xenopus caused by activation of MAP kinase.";
CC Nature 376:58-62(1995).
CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN
CC TYROSINE + PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: LOCALISED IN THE BRANCHIAL ARCH REGION AND
CC TAIL TIP.
CC -1- INDUCTION: BY RETINOIC ACID.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: CONTAINS THE CDC25 HOMOLOG DOMAINS 2 A AND B (CH2
CC DOMAINS A AND B).
DR EMBL: U43223; AAA85240.1; ...
DR HSSP: Q16828; IMP.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; Rhodanese_domain.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR SMART: SM00195; DSPC; 1.
DR SMART: SM00450; RHOD; 1.
DR PROSITE: PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR HydroLase.
FT DOMAIN 34 49 CH2 A DOMAIN.
FT DOMAIN 126 141 CH2 B DOMAIN.
FT DOMAIN 152 158 POLY-SER.
FT DOMAIN 205 378 CATALYTIC.
FT ACT_SITE 290 BY SIMILARITY.
SQ SEQUENCE 378 AA; 42036 MW; F274DACA960B783 CRC64;

Query Match 22.2%; Score 199; DB 13; Length 378;
Best Local Similarity 35.9%; Pred. No. 7.7e-13;
Matches 51; Conservative 23; Mismatches 38; Indels 30; Gaps 4;

QY 7 KILPGLYGNFDPARDADQLSKNKVTHLSV-----HDSPTL 43
DB 206 ELPLVLYGCAKDSYNLVLEEFKIKYILNVTPNLPLENAGEFRYKQIPISDHSONL 265

QY 44 CWTRHFKESIKFIHECRUGESC--LVHCLAGVRSYTLVAYIMTVTDFGMDALHTV 101
DB 266 ---SQFFPEAISFIDEA--RGKSGVLVHCLAGISRSYTVVAYIMOKLINSMDAYDIY 320

QY 102 RAGRCANPNVGFOROLOEFER 123
DB 321 KMKRSNISPNFNMGLDPER 342

RESULT 15
O99KC2 PRELIMINARY; PRT; 436 AA.
AC O99KC2:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIMILAR TO DUAL SPECIFICITY PHOSPHATASE 9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL: BC004738; AA04738.1; ...
SQ SEQUENCE 436 AA; 47088 MW; C4160F40DE263147 CRC64;

Query Match 22.2%; Score 199; DB 11; Length 436;
Best Local Similarity 35.0%; Pred. No. 9.1e-13;
Matches 49; Conservative 22; Mismatches 43; Indels 26; Gaps 2;

QY 7 KILPGLYGNFDPARDADQLSKNKVTHLSV-----HDSPTL 43
DB 258 QILPNLYLSARDSANLESIAKIGIRYILNVTPNLPLEKNDPFHKQIPISDHSONL 317

QY 44 CWTRHFKESIKFIHECRUGESC-LVHCLAGVRSYTLVAYIMTVTDFGMDALHTV 103
DB 318 ---SQFFPEAIAFIDEALSQNGCVLVHCLAGVRSYTVVAYIMOKLINSMDAYDLVR 374

QY 104 GRSCANPNVGFOROLOEFER 123
DB 375 KRSNISPNFNMGLDPER 394

Search completed: February 12, 2002, 16:16:27
Job time: 191 sec

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 12, 2002, 16:10:01 ; Search time 12.59 Seconds

(without alignments)  
298.495 Million cell updates/sec

Title: US-09-544-525-2

Perfect score: 897  
Sequence: 1 MGNMKNKILPGLYIGNFKDA.....AKNIIAAPGILKEMAFRLRL 167Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	186	20.7	397	2	US-08-990-379-8
2	169.5	18.9	207	4	US-09-013-881-4
3	168	18.7	353	4	US-09-013-881-3
4	167.5	18.7	226	4	US-09-045-973-8
5	161	17.9	367	2	US-08-530-290-24
6	159	17.7	393	2	US-08-990-379-4
7	159	17.7	394	2	US-08-530-290-23
8	159	17.7	395	2	US-08-990-379-5
9	158	17.6	263	4	US-09-164-193-5
10	156	17.4	313	2	US-08-990-379-7
11	156	17.4	314	2	US-08-990-379-5
12	152.5	17.0	198	4	US-09-163-833-2
13	152	16.9	367	2	US-08-990-379-6
14	150.5	16.8	198	4	US-09-045-973-5
15	147.5	16.4	185	1	US-07-988-273-2
16	147.5	16.4	185	3	US-08-848-810-25
17	147.5	16.4	185	4	US-09-164-193-21
18	147.5	16.4	185	5	PCT-US93-12019-2
19	143	15.9	72	2	US-08-530-290-20
20	138	15.4	117	1	US-07-988-273-4
21	138	15.4	117	5	PCT-US93-12019-4
22	138	15.4	117	5	PCT-US93-12019-19
23	121	13.5	72	2	US-08-530-290-15
24	109.5	12.2	116	1	US-07-988-273-5
25	109.5	12.2	116	5	PCT-US93-12019-5
26	99.5	11.1	118	5	US-07-988-273-3
27	99	11.0	201	2	PCT-US93-12019-3

28	99	11.0	201	4	US-09-234-613-21	Sequence 21, Appl
29	91.5	10.2	150	4	US-09-164-193-23	Sequence 23, Appl
30	86	9.6	209	4	US-09-164-193-8	Sequence 8, Appl
31	77.5	8.6	173	3	US-08-725-532A-3	Sequence 3, Appl
32	77	8.6	2465	2	US-08-596-291-3	Sequence 3, Appl
33	77	8.6	2465	3	US-09-100-804-3	Sequence 3, Appl
34	77	8.6	2466	3	US-09-080-855-12	Sequence 12, Appl
35	77	8.6	2466	5	PCT-US94-09943-2	Sequence 2, Appl
36	77	8.6	2485	4	US-09-230-640-46	Sequence 46, Appl
37	75.5	8.4	173	4	US-09-164-193-2	Sequence 2, Appl
38	75	8.4	212	1	US-08-461-859-35	Sequence 35, Appl
39	74	8.2	231	2	US-08-446-345-37	Sequence 37, Appl
40	74	8.2	242	2	US-08-685-992-21	Sequence 21, Appl
41	74	8.2	242	2	US-09-144-925-21	Sequence 21, Appl
42	73	8.1	45	2	US-08-530-290-21	Sequence 2, Appl
43	73	8.1	820	1	US-08-166-717D-6	Sequence 6, Appl
44	72	8.0	299	2	US-08-701-191A-13	Sequence 13, Appl
45	72	8.0	300	2	US-08-701-191A-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-08-990-379-8  
Sequence 8, Application US/08990379  
Patent No. 5998188  
GENERAL INFORMATION:  
APPLICANT: Stork, Philip J  
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and  
TITLE OF INVENTION: Their Biologically Active Expression Products  
FILE REFERENCE: 4104-000322USA  
CURRENT APPLICATION NUMBER: US/08/990.379  
CURRENT FILING DATE: 1997-12-15  
EARLIER APPLICATION NUMBER: PCT/US96/10402  
EARLIER FILING DATE: 1996-06-14  
EARLIER APPLICATION NUMBER: 60/000.263  
EARLIER FILING DATE: 1995-06-16  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 397  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-990-379-8  
Query Match 20.7%; Score 186; DB 2; Length 397;  
Best Local Similarity 33.6%; Pred. No. 1.5e-14;  
Matches 46; Conservative 21; Mismatches 52; Indels 18; Gaps 1;  
QY 4 GNNKILPGLYIGNFKDAEQDLSKNKYTHILSDPGLCWTR----- 48  
DB 177 GVEILPFLYLSAHSACEFLANLHTALLNVSRRSEACMTHLYKWIIVEDSHTD 236  
QY 49 ---HRESIKFTHECRKESCLVCLACVSRVTLVAVITVDFGMDALHTVRAGR 105  
DB 237 ISSHFOEALDFIDCKREKGVLVCHDEAGISSPTICMAYLMKTRQFRLEAFDIKQR 296  
QY 106 SCANNVGFQROLOEFE 122  
DB 297 SWVSPNFGFMGLQYE 313  
RESULT 2  
US-09-013-881-4  
Sequence 4, Application US/09013881  
Patent No. 6132964  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Hillman, Jennifer L.

```

Query Match          18.9%; Score 169.5; DB 4; Length 207;
Best Local Similarity 31.0%; Pred. No. 6.4e-13;
Matches 48; Conservative 24; Mismatches 52; Indels 31; Gaps 5;

QY      3  NGMKNIKLEGLYIGNFKDARDAEOLSKNKVTHLLSVHDSFGLCWR-----46
      |:::||||:|:  |:::||||:|  |
Db      60  NHADEWMPGLTIGDQDMANNRRLEIRLIGITVHLNASHSR---WRGTPEAYESTGIRLGV 116
      |:::||||:|:  |:::||||:|  |

QY      47  -----TRHFESIKLFTHECRRLR-GESELYVACLAGSVSYVLVYMTVYDFGEMDAIH 99
      |:::||||:|:  |:::||||:|  |
Db      117 EPADFMSITHFOTADFIHRLASOPGKLLVHCAGVSRATLVLYMLVHNHLLVLEAIK 176
      |:::||||:|:  |:::||||:|  |

QY      100 TVRAGSCANPNVGFOROLOEFEEKHEVHYQYOMWLK 134
      |:::||||:|:  |:::||||:|  |
Db      177 KVKDHRGII-PNRGFLRQLALDR---RLRQGLE 206
      |:::||||:|:  |:::||||:|  |

RESULT      3
US-09-013-881-3
; Sequence 3, Application US/09013881
; Patent No. 6132964

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Shah, Purvi

TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES

```

```

Query Match      18.7%; Score 168; DB 4; Length 353;
Best Local Similarity 29.3%; Pred. No. 2,le-12;
Matches 49; Conservative 27; Mismatches 59; Indels 32; Gaps 6;

OY      5 MNKILPGLYIGNEPKDARDAEQLSKNKVTHILSVHD-----SPGL--CMR----- 46
           | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      40 MLEVOPGLYFPGCAAAVAAPEDHLREAGITAVLTVDSEEPSFKAGPEDEMLRFLFVPALDKP 99
OY      47 -----TRHKESIKITFHEERLGEISCLVCHGVSRYTLVATVITVDFGEMDLHYR 102
           | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      100 ETDLLSHLDRCVAFILGAKRAAGRAVLVCHGVSRSVAITVAFLEKTKDQLPEKAYEKIQ 159
OY      103 AGRSCANPNVGFQROLOEFER--HEV-----HOYR--OMLKEEYGE 139
           | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      160 ILKPEAKMNEGEFEMWLKLYQAMGYEVDTSIAIKYQRIQLQKTEKXPE 206

RESULT      4
US-09-045-973-8
; Sequence 8, Application US/09045973
; Patent No. 6165767
;
; GENERAL INFORMATION:
;
; APPLICANT: Lal, Preeti
;
; APPLICANT: Yue, Henry
;
; APPLICANT: Corley, Neil C.
;
; APPLICANT: Guejler, Karl J.
;
; APPLICANT: Baughn, Mariah
;
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
;
; NUMBER OF SEQUENCES: 9
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;
; STREET: 3174 Porter Drive
;
; CITY: Palo Alto

```

STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/045,973  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0491 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1495338  
US-09-045-973-8

Query Match 18.7%; Score 167.5; DB 4; Length 226;  
Best Local Similarity 28.6%; Pred. No. 1.3e-12;  
Matches 48; Conservative 33; Mismatches 58; Indels 29; Gaps 5;  
QY 5 MNKILPGIYGNFKARAEQLSKNKYTHILSV-----SPGLCKRTR----- 48  
DB 14 MSEIIPGLFICGV-SALSKEDEKKRKHITHIINATTEVPNLSIDIOPTKMLIEDTPQTY 72  
QY 49 ---HKESIKFHECRLEGESCLVHCLAGSVSTLVATYMTYVDFGEMDALHTVRAGR 105  
DB 73 IYPIELIOSDIOIALIADGKVLVHCVAGSVRSASICIAFLIKYRCRMUREYVHMKSR 132  
QY 106 SCANPNVGFQROLOEFKEHEVHQYRWMLKEEYESPL-QDAEAKNIL 152  
DB 133 SMVRNGLFWQLIAYEON-----VKENAGSVRLVRDEAQPOLL 172

RESULT 5  
US-08-530-290-24  
Sequence 24, Application US/08530290  
Patent No. 5958721  
GENERAL INFORMATION:  
APPLICANT: Marshall, Christopher John  
APPLICANT: Ashworth, Alan  
APPLICANT: Hughes, David Anthony  
TITLE OF INVENTION: Methods for Screening of Substances for  
TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/530,290  
FILING DATE: 14-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/GB94/00694  
FILING DATE: 31-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9402573.1  
FILING DATE: 10-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9307250.2  
FILING DATE: 07-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 084611-00000005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 367 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-530-290-24

Query Match 17.9%; Score 161; DB 2; Length 367;  
Best Local Similarity 29.8%; Pred. No. 1.6e-11;  
Matches 42; Conservative 20; Mismatches 53; Indels 26; Gaps 2;  
QY 4 GWNKILPGIYGNFKARDARQISKNKYTHILSV-----HDSF 41  
DB 173 GPVEILPFLVIGSAVHASKRMDLALGITALLINVSANCPNHFEGHYQKSIPEVDNHRKD 232  
QY 42 GICWTRHFKESIKFHECRLEGESCLVHCLAGSVSTLVATYMTYVDFGEMDALHTV 101  
DB 233 ISSW----FNEAIDFIDISIKNAGRGVFWHCQGISRSATICLAYIMRNRYKIDEAFERY 288  
QY 102 RAGRCANPNVGFQROLOEFE 122  
DB 289 KORRSIIISPNSEFMGOLLQFE 309

RESULT 6  
US-08-990-379-4  
Sequence 4, Application US/08990379  
Patent No. 5998188  
GENERAL INFORMATION:  
APPLICANT: Stock, Phillip J  
APPLICANT: Mista-Press, Anita  
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase CDNs and  
TITLE OF INVENTION: Their Biologically Active Expression Products  
FILE REFERENCE: 4104-000322USA  
CURRENT APPLICATION NUMBER: US/08/990,379  
EARLIER FILING DATE: 1997-12-15  
EARLIER APPLICATION NUMBER: PCT/US96/10402  
EARLIER FILING DATE: 1996-06-14  
EARLIER APPLICATION NUMBER: 60/000,263  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 393  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-08-990-379-4

[illegible]



EARLIER APPLICATION NUMBER: PCT/US96/10402  
EARLIER FILING DATE: 1996-06-14  
EARLIER APPLICATION NUMBER: 60/000,263  
EARLIER FILING DATE: 1995-06-16  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO: 6  
LENGTH: 367  
TYPE: PRT  
ORGANISM: Mus sp.  
US-08-990-379-6

Query Match 16.9%; Score 152; DB 2; Length 367;  
Best Local Similarity 29.1%; Pred. No. 2.1e-10;  
Matches 41; Conservative 20; Mismatches 54; Indels 26; Gaps 2;

QY 4 GNMKILPGIYICNFKDAEQLSKNKVTHILSV-----HDSF 41  
DB 173 GPEIILSFLIGSAYHSRDMDLALGITALINVSANCPNHFGHYQKSIPEVDNKKAD 232  
QY 42 GLCWRTHFESIKFHECRLEGSCLYHCLAGVSRYLVYIMTVDFGMDALHTY 101  
DB 233 ISSW-----FNEAIDFIDISIDAGRVFVHCAGISRSATTCIAYIMTRNKVKLDEAFEFV 288  
QY 102 RAGSCANPNVGFQROLOEFE 122  
DB 289 KORRSIISPNFSFMGQLOPE 309

RESULT 14  
US-09-045-973-5  
Sequence 5, Application US/09045973  
Patent No. 6165767  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Tue, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Guebler, Karl J.  
APPLICANT: Baughn, Mariah  
TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/045,973  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PR-0491 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRSTNOT16  
CLONE: 3041794  
US-09-045-973-5

Query Match 16.8%; Score 150.5; DB 4; Length 198;  
Best Local Similarity 27.7%; Pred. No. 1.3e-10;  
Matches 48; Conservative 30; Mismatches 76; Indels 19; Gaps 3;

QY 4 GNMKILPGIYICNFKDAEQLSKNKVTHILSVH-DSPGLCWRT----- 47  
DB 26 GIAQITSSFLIGSAYHSRDLQARGITCIYVATTEIPNFMPQFEYKVPYLDMPHAP 85  
QY 48 --RHFKESIKFHECRLEGSCLYHCLAGVSRYLVYIMTVDFGMDALHTYRGR 105  
DB 86 IGLYEDTVADKTHSVSRKIGATLVHCAAGVSRSATTCIAYIMKFNVCILAEYMWVKAR 145  
QY 106 SCANPNVGFQROLOEFEKHEVHQR-OWLKEEYGESEPLQDAEAKNIIAARPI 157  
DB 146 PYIRPNVGFWRQIIDYEROLFEGKSTVKWVQPPYGLVPPVYEKSHLMPYGI 198

RESULT 15  
US-07-988-273-2  
Sequence 2, Application US/07988273  
Patent No. 5512434  
GENERAL INFORMATION:  
APPLICANT: AARONSON, Stuart A.  
APPLICANT: BOTTARO, Donald P.  
APPLICANT: ISHIBASHI, Toshio  
APPLICANT: MIKI, Toru  
TITLE OF INVENTION: EXPRESSION CLONING OF A HUMAN  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington, D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/988,273  
FILING DATE: 19921214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40399/182 NIHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 185 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-988-273-2

Query Match 16.4%; Score 147.5; DB 1; Length 185;  
Best Local Similarity 29.8%; Pred. No. 2.8e-10;  
Matches 42; Conservative 20; Mismatches 50; Indels 29; Gaps 4;









BASE COUNT 233 a 230 c 242 g 221 t  
ORIGIN

Query Match 91.7%; Score 802.4; DB 6; Length 926;  
Best Local Similarity 94.4%; Pred. No. 1.4e-200;  
Matches 874; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

```
OY 1 cccgcgcgcctccctccctgttaacatgcatagtgctgtgagccacacgcccggc 60
Db 1 cccgcgcgcctccctccctgttaacatgcatagtgctgtgagccacacgcccggc 60
OY 61 gctagcgtctgcctccacacacatggggaatggaatgaacaaagatcccgccgtgt 120
Db 61 gctagcgtctgcctccacacacatggggaatggaatgaacaaagatcccgccgtgt 120
OY 121 acatgggaacttcaaatgacagagagacgagaaacattgagcaagaagaagtgacac 180
Db 121 acatgggaacttcaaatgacagagagacgagaaacattgagcaagaagaagtgacac 180
OY 181 atattctgtctgcacagataagt-ccaggcctatgtttgag----- 220
Db 181 atattctgtctgcacagataagt-ccaggcctatgtttgag----- 220
OY 221 -----gacaagaacattcaagaagaattataat 249
Db 241 TCCACAGACGGAGATTTCACATCTCAAAACCTGACAAAGACATTTCAAAGAAAGTATTAAAT 300
OY 250 tcaattacagagtgccggtctccggtgagagctgctgttaacatgctgtgcccgggtct 309
Db 301 TCATTTCAGAGAGTCCGGCTCCGGGTGAGAGTCCCTTGACACTGCTGGCCGGGTCT 360
OY 310 ccaggagcgagaaactgagatgcatacatgacacgctacatgctgttgggtgagag 369
Db 361 CCAGAGAGCGTACACTGAGTATGACATCTGACCTGACACTGCTGTGGTGGAGG 420
OY 370 atgcccctgcaacacgtgctgcccggagagatcctgtgccaaccccaacgtgagcttcaca 429
Db 421 ATGCTCTGACACCGTGTGCTGTGGAGATCTCTGCAACCCCAACGTTGGCTTCACA 480
OY 430 gacagctcagagagtttgagagacatgagatgcatacagatgcagagtggttgaagaag 489
Db 481 GACAGCTCCAGAGATTGAGAACATGAGGTCCATCATGATGCGCAGTGGCTGAAGAGAG 540
OY 490 aatattgagagagcccttgcagagatgcagaaagacaaacacattcggccggtccag 549
Db 541 AATATGAGAGAGCCCTTTGCAAGATGACGAAGAACCCAAAACATTTCTGGCCGCTCCAG 600
OY 550 gaattctgaagttctgggcttctccaagaagactglaattgactgaaagtttctgaata 609
Db 601 GAATTCTGAAGTTCTGGGCTTTCTCAGAACTGTAATGTAACGTGAAGTTCTGAATA 660
OY 610 ttgcaaacccgcagagtttaggctgtgtgctgccaagaagaagaacacacatagatttaag 669
Db 661 TTGCAAAACCCGAGAGTTTAGGCTGTGCTGCCAAAAGAAAAGCAACATAAGATTAAAG 720
OY 670 tatccagtagtgatttgtaaacattgttttcaattgaaagctgaacatactagtagtag 729
Db 721 TATTCAGTAGTAGTATTGTAACCTGTGTTTTCATTGGAAGCTGAATATATACGTATC 780
OY 730 tttatgttgagagaactcaagagatactctttagcaagaagaataatttccctatccac 789
Db 781 TTTATGTTGAGAACTTAAGATATCTTTAGCAAGAGAAAATATTTTCCCTTATCCAC 840
OY 790 tgcctgtagaggttctgtacccgttgatgagcctgtaagatcccgggaacccctgccc 849
Db 841 TGCTGTGAGAGTTTCTGTAACCTCGCTTGATGCTCGTAAGATCCCGGACCTTGGCCG 900
OY 850 actgacctgtggtgagcttgagcgtc 875
Db 901 ACTGCTTGTGGGTGGCTTGGGGCTC 926
```

RESULT 2

AX068328

LOCUS

AX068328

DEFINITION

AX068328

ACCESSION

AX068328

VERSION

AX068328.1

KEYWORDS

GI:12578509

SOURCE

human.

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 926)

AUTHORS

Luche, R.M. and Wei, B.

TITLE

DSP-3 dual-specificity phosphatase

JOURNAL

Patent: WO 0102582-A 1 11-JAN-2001;

Ceplyr, Inc. (US)

FEATURES

Location/Qualifiers

1..926

source

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 233 a 230 c 242 g 221 t

ORIGIN

Query Match 91.7%; Score 802.4; DB 6; Length 926;  
Best Local Similarity 94.4%; Pred. No. 1.4e-200;  
Matches 874; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

```
OY 1 cccgcgcgcctccctccctgttaacatgcatagtgctgtgagccacacgcccggc 60
Db 1 cccgcgcgcctccctccctgttaacatgcatagtgctgtgagccacacgcccggc 60
OY 61 gctagcgtctgcctccacacacatggggaatggaatgaacaaagatcccgccgtgt 120
Db 61 gctagcgtctgcctccacacacatggggaatggaatgaacaaagatcccgccgtgt 120
OY 121 acatgggaacttcaaatgacagagagacgagaaacattgagcaagaagaagtgacac 180
Db 121 acatgggaacttcaaatgacagagagacgagaaacattgagcaagaagaagtgacac 180
OY 181 atattctgtctgcacagataagt-ccaggcctatgtttgag----- 220
Db 181 atattctgtctgcacagataagt-ccaggcctatgtttgag----- 220
OY 221 -----gacaagaacattcaagaagaattataat 249
Db 241 TCCACAGACGGAGATTTCACCATCTCAAAACCTGACAAAGACATTTCAAAGAAAGTATTAAAT 300
OY 250 tcaattacagagtgccggtctccggtgagagctgctgttaacatgctgtgcccgggtct 309
Db 301 TCATTTCAGAGAGTCCGGCTCCGGGTGAGAGTCCCTTGACACTGCTGGCCGGGTCT 360
OY 310 ccaggagcgtagaaccttgtagatgcatacatgacacgctacatgctgttgggtgagag 369
Db 361 CCAGAGAGCGTACACTGAGTATGACATCTGACCTGACACTGCTGTGGTGGAGG 420
OY 370 atgcccctgcaacacgtgctgcccggagagatcctgtgccaaccccaacgttgaaga 429
Db 421 ATGCTCTGACACCGTGTGCTGTGGAGATCTCTGCAACCCCAACGTTGGCTTCACA 480
OY 430 gacagctcagagagtttgagagacatgagatgcatacagatgcagatgctgaagaag 489
Db 481 GACAGCTCCAGAGATTTCAGAGATGAGGTCCATCATGATGCGCAGTGGCTGAAGAG 540
OY 490 aatatgagagagcccttgcagagatgcagaagaagacaaacacattcggccgtccag 549
Db 541 AATATGAGAGAGCCCTTTGAGAGATGACGAAGAACCCAAAACATTTCTGGCCGCTCCAG 600
OY 550 gaattctgaagttctgggcttctccaagaagactglaattgactgaaagtttctgaata 609
Db 601 GAATTCTGAAGTTCTGGGCTTTCTCAGAACTGTAATGTAACGTGAAGTTCTGAATA 660
OY 610 ttgcaaacccgcagagtttaggctgtgtgctgccaagaagaagaagcagacataagatttaag 669
```

```

Db 661 TTGAAACCGCAGAGTTTAGCGCTGGTCCCAAAAGAAAGCAACATAGACTTTAG 720
OY 670 tatcagtagtattgttaactgttttcaatttgaagcgaalatalacgtatcatg 729
Db 721 TATCAGTAGTGAATTTTAAACTGTTTTCATTGTAAGCTGATATATAGTAGTCATG 780
OY 730 ttatgtgtgaactaagaatattctttagcaagaagaataatttccctatccccac 789
Db 781 TTTATGTGAGAACTAAGATATTTCTTACCAAGAGAAATATTTCCCTTATCCCCAC 840
OY 790 tgcgtggaagtttcttactcgtcgtgagatgcctgtaagaatccggagaccttgcgc 849
Db 841 TGCgtgagagtttcttactcgtcgtgagatgcctgtaagaatccggagaccttgcgc 900
OY 850 actgacctgtggtgtgcttgcgc 875
Db 901 ACTGCCCTTGTGGGTGGCTTGGCGCTC 926

RESULT 3
AX086014 AX086014 1379 bp DNA PAT 09-MAR-2001
LOCUS Sequence 11 from Patent WO0112819.
ACCESSION AX086014
VERSION AX086014.1 GI:13275838
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1379)
AUTHORS Ploemhan, G.D., Martinez, R., Whyte, D., Hill, R., Flanagan, P. and
Joubin, M.
TITLE Protein phosphatases and diagnosis and treatment of
phosphatase-related disorders
JOURNAL Patent: WO 0112819-A 11 22-FEB-2001;
Sugen, Inc. (US)
FEATURES
Location/Qualifiers
Source 1. 1379
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 298 a 374 c 386 g 320 t 1 others
ORIGIN
Query Match 91.3%; Score 799.2; DB 6; Length 1379;
Best Local Similarity 94.2%; Pred. No. 1e-199;
Matches 872; Conservative 0; Mismatches 3; Indels 51; Gaps 2;

OY 1 cccgcgcgtctctctctcctcctgttaacatgcatatgtgcgtgcgaacacgcgcgcgcgc 60
Db 257 CCCGCGCGCTCCCTCCCTCCCTGTAACATGCGCATAGTGGCCCGCACGACGCGCGGCGC 316
OY 61 gctagcgttcgccttcgacccacacacacacacacacacacacacacacacacacacacac 120
Db 317 GCTAGCGTTCCTTCACCCACCATGGGAAATGGAGTAACAAGATCTCTGCGCGCGCTGT 376
OY 121 acatcgcaacttcaaaagatccagagacacacacacacacacacacacacacacacacac 180
Db 377 ACATCGGCAACTTCAAGATGCCAGAGCGGGAACAATTAGCAAGAACAGAGTGAGACAC 436
OY 181 atattcgtctgcacagatgt-ccagagcattgttggag----- 220
Db 437 AATATTCGTCTGTCCACAGATAGTCCAGGCTATGTTGAGAGGAGTAAATACCTGTGCA 496
OY 221 -----gacaagacatttaaaagaagaatttaaat 249
Db 497 TCCACAGACGGGANTTACCATTCTCAAAACCTGACACAGACATTTTCAAAAGATATTAAAT 556
OY 250 tcatcacaagatgcgcgcgtcgcgcgtgtgagagctgcctgttaacatgcctgcgcgcgcgc 309
Db 557 TCATTTCAGAGATGCCGCTCCGCGGTGAGAGCTGCTTGTACACTGCTTGGCGCGGCGTCT 616

```

```

OY 310 ccaagagcgtgacactgtgtgacatcacatcacatgacgcgtactgactgttgcgcgcgcgc 369
Db 617 CCAGAGACGCTGACACTGCTGATGATCATATCATGACACCGTCACTGACTTGTGGGAGG 676
OY 370 atgcccgcgaacacgcgtgcgtccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 429
Db 677 ATGCCCTGACACCCGCTGCTGCTGAGATCTCTGTGCAACCCACACCTGGGCTTCCAGA 736
OY 430 gacagctccaaagatttgaagaacatgagtcacatcagatccgcgcgcgcgcgcgcgcgcgc 489
Db 737 GACAGCTCCAGAGCTTGTGAGAACGATGAGTGCATCATATGCGCATGCGGTGAAGAGAG 796
OY 490 aatctggaagaagcccttgcagagatgcagaagaagccaaacacacacacacacacacac 549
Db 797 AATATGAGAGAGCCCTTTCAGAGATGACAGAAAGCCAAACACATTCCTGGCCGCTCCGG 856
OY 550 gaattctgaagttctgggccttctcagaagaacgtatgtactcgaagtttctgaata 609
Db 857 GAATTCGAAAGTTCTGGGCTTCTCTCAGAAACGTGTATGTACCTGAAGTTCTGAAATA 916
OY 610 ttgcaaacccgcagagtttagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 669
Db 917 TTGCAAAACCCACAGAGTTTAGGCTGTGCTGCTGCCAAAAGAAAGCAACATAGATTAG 976
OY 670 tatcagtagtattgttaactgttttcaatttgaagcgaatatalacgtatcatg 729
Db 977 TATCCAGTAGTGAATTTGAAACTGTTTTCATTGTAAGCTGATATATATAGTATCATMG 1036
OY 730 ttatgtgtgaactaagaatattctttagcaagaagaataatttccctatccccac 789
Db 1037 TTTATGTGAGAACTAAGATATTTCTTACGAAAGAGAAATATTTCCCTTATCCCCAC 1096
OY 790 tgcgtggaagtttcttactcgtcgttgaagcgtgtgaagatcccgagagccttgcgc 849
Db 1097 TGCgtgagagtttcttactcgtcgttgaagcgtgtgaagatcccgagagccttgcgc 1156
OY 850 actgacctgtggtgtgcttgcgc 875
Db 1157 ACTGCCCTTGTGGGTGGCTTGGCGCTC 1182

RESULT 4
AF165519
LOCUS AF165519 1092 bp mRNA PRI 20-JUL-2000
DEFINITION Homo sapiens mitogen-activated protein kinase phosphatase x (MKPX)
ACCESSION AF165519
VERSION AF165519.1 GI:9294744
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1092)
AUTHORS Gu, J., Huang, Q., Yu, Y., Xu, S., Wang, Y., Han, Z., Zhou, J.,
Tu, Y., Gu, W., Fu, G. and Huang, C.
TITLE Novel genes expressed in hematopoietic stem/progenitor cells from
Myelodysplastic Syndromes patient
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1092)
AUTHORS Gu, J., Huang, Q., Yu, Y., Xu, S., Wang, Y., Han, Z. and Chen, Z.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
Source 1. 1092
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="hematopoietic stem/progenitor cell"
1. 1092
/gene="MKPX"

```



Db 549 GCGGCCCGGGAATTCGAGTACTGGGCTTCTCGAAGACGTGATGTACTGAGT 608  
QY 600 ttctgaataatgcaaacccgcagagtttagctgtgtgtcgcacaaagaagaacacat 659  
Db 609 TTCTGAATATTCGCA-----AGTTCAAGCTGTGCTGCCAAGAAAGAGTGTATGT 660  
QY 660 aagg-----ttaagatccagtagtgatttgaacttg-----ttttcatttgaagc 709  
Db 661 AAAGTTTATTTTAAAGCAATAGATGATTTGTATCTGTTTTCATTTCATTTCATTAAC 720  
QY 710 tgatataatagtagtcaatgtt-----taagtgaactaagatattcttgaagaag 765  
Db 721 CAATGCAATGATATATATGTGTGGAATATGTGATATCTTGTGATATTCGTAGCAGAG 780  
QY 766 aaaaatttccctatcccccgcgtcgtgaggttct 805  
Db 781 AAAATATCTTGGCTTAACCTCACTGCTGTGTTGCT 820

RESULT 6  
AF237619 1067 bp mRNA ROD 17-JUL-2001  
LOCUS AF237619 1067 bp mRNA ROD 17-JUL-2001  
DEFINITION Mus musculus dual specificity phosphatase TS-DSP2 mRNA, complete cds.  
ACCESSION AF237619  
VERSION AF237619.1 GI:13183068  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 1067)  
AUTHORS Aoyama,K., Nagata,M., Oshima,K., Matsuda,T. and Aoki,N.  
TITLE Molecular cloning and characterization of a novel dual specificity phosphatase, LMR-DSP2, that lacks the Cdc25 homology domain  
JOURNAL J. Biol. Chem. 276 (29), 27575-27583 (2001)  
MEDLINE 21347909  
PUBMED 11346645  
REFERENCE 2 (bases 1 to 1067)  
AUTHORS Aoyama,K., Matsuda,T. and Aoki,N.  
TITLE Direct Submission  
JOURNAL Submitted (19-FEB-2000) Applied Molecular Biosciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan  
FEATURES  
source  
1.1067  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
28..582  
CDS  
/codon\_start=1  
/product="dual specificity phosphatase TS-DSP2"  
/protein\_id="AAK15038.1"  
/db\_xref="GI:13183069"  
/translation="MGSCMSQILPGLYIGNFKDAPAEQLSRNKHVHLSVHDTPARM  
LEGVYKICPADPTPSONLTRHKESIKRFLGEGSLVGLAGVSRVLTIVAY  
IKMTVDEGMDALHYRAGRSKANPNLGRORLOEFKHEVHOYRQMLREYGENPLR  
DAEFKNIITAAPEITLYMAFLRL"  
BASE COUNT 301 a 223 c 260 g 283 t  
ORIGIN

Query Match 51.28; Score 447.8; DB 10; Length 1067;  
Best Local Similarity 79.88; Pred. No. 3.6e-107;  
Matches 633; Conservative 0; Mismatches 87; Indels 73; Gaps 6;

QY 78 gccaccatggaatggatgaacaagatcctgcggcgccctgtcatcgcgaacttcaaa 137  
Db 22 GCGGCCCATGGGAGTGGGATGAGCCAGATCTGCGGCGCTGTACATTTGGCAACTTCAAA 81  
QY 138 gatgcagagacgcgaacaattgagcaagaaggtagacacatattctgtcgtcac 197  
Db 82 GACGCAAGAGATGCAGAAAGTTGAGCAGAAAGGTGACACACATTCTTTCTGTGCAC 141

QY 198 gatagt-ccaagcctatgttggag----- 220  
Db 142 GATACCTGCCAGGCCCATGTTTGAGGAGATTAAATACCTGTGTATTTCACAGCGGACAGACA 201  
QY 221 -----gacaagacattccaagaagatlaaattcattcagagtgccgg 266  
Db 202 CCATCTCAAAACCTGACAGACATTTCAAGAAAGCATTTAATTCATTTCATGATGCGGA 261  
QY 267 ctccgcgtgagagctgtccttgaacactgcctgcgggggtctccaggaagctgaacatg 326  
Db 262 CTCAGGGGTGAGAGCTGTCTGTACATTTGCTGCTGCGGGGCTCCAGAGAGTGCATTTG 321  
QY 327 gtagatcatatcatcagaccgttactgacttggctcgggagagatgcccctgacacgctg 386  
Db 322 GTGATCCCATACATCATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 381  
QY 387 cgtgcgggaatcctctgtccaccccaagctggagcttccagagacagctccagagatt 446  
Db 382 GGTGCGGGAGAGTCTGTGCTGCAACCCCAACCTGGGCTTTCAAAAGGACGCTCAGGAGTTT 441  
QY 447 gagaagcatgagtcacatcagatcagtcagtcgagtcgaagaagaatataatgagagaccct 506  
Db 442 GAGAACATGAGAGTGCACACCATATCGGCAATGCGCTGAGAGAGATGAGAGAGACCT 501  
QY 507 ttgcagatgcagaagaagccaaacacatctggccgctccaggaatttgaagttctgg 566  
Db 502 TTGCGGATGTCAGAGAAAGCCAAAAATTTCTGCTGCGCCGCGGAAATTCGTAAGTACGG 561  
QY 567 gcccttctcgaagaacgttaattgtactcgaagtttctgaatatattgcaaccgcagagt 626  
Db 562 GCCCTTCTCAAAACATGTAATGTACGAGGTTCTGAAATATTCGCA-----AGT 613  
QY 627 ttagctgtgtgtccaaagaagaagaacatagag-----ttaagtctcagtagt 680  
Db 614 TTAGCTGTGTGTCTGCCAAAAAGAAAGATGATTAAGTTTATTTTAAGATCCATATGT 673  
QY 681 gatttgaacttg-----ttttcatttgaactgaatatacagtagtcatgtt---t 732  
Db 674 GATTGTATATCTGTTTTCATTTTTCATTTTAAACCAATGATATATGATGTTGAAT 733  
QY 733 atgttgaagactaaggatattctttagcaagaagaatatltccctatcccatgctgc 792  
Db 734 ATGTGAAGATCTATGATATCTGTAGCAGAAAGAAATCTTTCCTTAATCTCAGTGC 793  
QY 793 tgtgaggttct 805  
Db 794 TGTGTTGTTCT 806

RESULT 7  
AX068352  
LOCUS AX068352 555 bp DNA PAT 25-JAN-2001  
DEFINITION Sequence 25 from Patent WO0102582.  
ACCESSION AX068352  
VERSION AX068352.1 GI:12578520  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 555)  
AUTHORS Luche,R.W. and Wei,B.  
TITLE Dsp-3 dual-specificity phosphatase  
JOURNAL Patent: WO 0102582-A 25 JAN-2001;  
Ceptyr, Inc. (US)  
FEATURES  
source  
1..555  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 149 a 136 c 154 g 116 t  
ORIGIN

Query Match 49.3%; Score 431.4; DB 6; Length 555;  
 Best Local Similarity 90.6%; Pred. No. 7.4e-103;  
 Matches 503; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

QY	84	atgggggaatggatgaacaagatcctgcggccctgtacatcggcgaactcaagaatgc	143
DB	1	ATGGGGAATGGATGAAACAGATCCGCCGCCCTGTACATCGGCAACTCAAAAGTGGC	60
QY	144	agaaagcggaacaatttgcaagaacaggtgacacatatctgtctgtccacgaagt	203
DB	61	AGAGAGCGCGAACAATTGAGCAAGAACAGGTGACACATATTTCTGTCTCCACGATAGT	120
QY	204	-ccaaagcctatgttgag-----	220
DB	121	GCAGAGCCTATGTTGGAGGAGTTAAATACCTGTGTGATCCAGACGGATTACCATCT	180
QY	221	-----gacaagaattccaagaagaatataatcattcacaagatgcccgtccgc	272
DB	181	CAAAACCTGACAAAGACATTTCAAGAAATTAATTATTCACAGATGCCGGCTCCGC	240
QY	273	agtgagaagctgctgtacacatgcctgcgggggtctccaggagcgtgacactgtatc	332
DB	241	GGTGAAGAGTGCCTTGTACACTGCTGGCCGGGGTCTCCAGAGCGTGTACACTGGTATC	300
QY	333	gcatacatcatgaccgtcactgacttggctgtggaagatgcctgcacacgltgcgtcc	392
DB	301	GCATACATCATGACCGTCACTGACTTTGGCTGGAGGATGCCCTGACACCGTGGCTGCT	360
QY	393	gggaagatcctgtgccaaccccaagctgggtctccagaagacagctccaggagttgaag	452
DB	361	GGGAATCTCTGTGCCAACCCCAACGGGGCTTCCAGAGACACTCCAGAGTTTGAAG	420
QY	453	catgagtcacatcatatcgtcagctgagctggaagaagaatgatggaagacccctttgcg	512
DB	421	CATGAGTCTCATCATGATATGGCATGCTGGAAGAAATATGAGAGAGCCCTTTGCG	480
QY	513	gatgcagaagaagccaacaacattctgacgctccaggaaattcgaagttctgacgtt	572
DB	481	GATGCAAGAAAGAACAAAACATCTGGCGCTCCAGGAATTTCTGAAGTTCTGGGCTTT	540
QY	573	ctcagaagaagcttaa 587	
DB	541	CTCAGAAGACTGTAA 555	

RESULT 8  
 LOCUS AL365272 132449 bp DNA PRI 23-DEC-2000  
 DEFINITION Human DNA sequence from clone RP11-328C17 on chromosome 6, complete sequence.  
 ACCESSION AL365272  
 VERSION AL365272.20 GI:11991402  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 132449)  
 AUTHORS Heath, P.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 On Dec 24, 2000 this sequence version replaced gi:11875910.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence has been finished according to sequence map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known

annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone configs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> RP11-328C17 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-328C17 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone RP11-328C17 is at 132449 in this sequence. The true left end of clone RP11-233K4 is at 64893 in this sequence. The true right end of clone RP3-416J7 is at 100 in this sequence. The true right end of clone RP1-125A24 is at 61243 in this sequence.

## FEATURES

source	Location/Qualifiers
	1..132449
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="6"
	/clone="RP11-328C17"
	/clone_11b="RP11-11.2"
	1..1562
repeat_region	/note="LIM4 repeat: matches 4133..5719 of consensus"
	1563..1641
repeat_region	/note="MLT1A2 repeat: matches 1..80 of consensus"
	1659..2199
repeat_region	/note="LIM4 repeat: matches 3396..3937 of consensus"
	2200..2474
repeat_region	/note="Alusx repeat: matches 38..312 of consensus"
	2475..4279
repeat_region	/note="LIM4 repeat: matches 1686..3396 of consensus"
	4306..4589
repeat_region	/note="LIM4 repeat: matches 1254..1544 of consensus"
	complement(4704..5324)
misc_feature	/note="match: GSS: Em:AQ628183"
	4799..4938
repeat_region	/note="LIM4c repeat: matches 785..924 of consensus"
	complement(4809..5326)
misc_feature	/note="match: GSS: Em:AQ774852"
	complement(4860..5272)
misc_feature	/note="match: GSS: Em:AQ787689"
	4943..5094
repeat_region	/note="LIM4c repeat: matches 272..420 of consensus"
	complement(5020..5329)
misc_feature	/note="match: GSS: Em:AQ494601"
	5142..5435
repeat_region	/note="MLT1D repeat: matches 170..428 of consensus"
	5436..5744
repeat_region	/note="Alusx repeat: matches 1..310 of consensus"
	5745..5869
repeat_region	/note="MLT1D repeat: matches 428..568 of consensus"
	6450..6686
repeat_region	/note="MIR repeat: matches 7..261 of consensus"
	6916..7994
repeat_region	/note="LIM4 repeat: matches 3582..4658 of consensus"
	7984..8050
repeat_region	/note="LIMB8 repeat: matches 5808..5884 of consensus"
	8064..8169
repeat_region	/note="LIM4 repeat: matches 3479..3593 of consensus"
	8178..8224
repeat_region	/note="LIMB4 repeat: matches 6124..6169 of consensus"
	8225..8525
repeat_region	/note="Alusx repeat: matches 1..301 of consensus"
	8526..8915

```
misc_feature /note="LIMB4 repeat: matches 5725. .6124 of consensus"
complement(8862. .9353)
repeat_region /note="match: GSS: Em:AQ355497"
9029. .9150
/note="2 copies 61 mer 95% conserved"
misc_feature complement(9052. .9484)
/note="match: GSS: Em:AQ175845"
9388. 9675
/note="match: GSS: Em:AQ135206"
misc_feature 9487. .9840
/note="match: GSS: Em:AQ171684"
9491. .9932
misc_feature /note="match: GSS: Em:AQ147759"
9524. .9697
/note="MER33 repeat: matches 1. .189 of consensus"
repeat_region 10046. .10143
/note="2 copies 49 mer 91% conserved"
repeat_region 11168. .11361
/note="MTR18 repeat: matches 101. .309 of consensus"
11448. .11756
repeat_region /note="AluY repeat: matches 1. .309 of consensus"
12109. .12515
repeat_region /note="L2 repeat: matches 1747. .2155 of consensus"
12798. .12947
repeat_region /note="MIR repeat: matches 29. .206 of consensus"
12975. .13026
repeat_region /note="MIR repeat: matches 186. .237 of consensus"
13005. .13103
/note="L2 repeat: matches 2417. .2516 of consensus"
repeat_region 13331. .13664
/note="MTR18 repeat: matches 4. .361 of consensus"
13704. .13907
repeat_region /note="LRR33 repeat: matches 297. .511 of consensus"
15450. .15530
repeat_region /note="L2 repeat: matches 2668. .2749 of consensus"
15876. .15935
repeat_region /note="2 copies 30 mer 100% conserved"
16202. .16294
repeat_region /note="MER81 repeat: matches 2. .94 of consensus"
17335. .17668
misc_feature /note="match: GSS: Em:AQ104665"
17419. .17715
repeat_region /note="AluX repeat: matches 1. .296 of consensus"
17995. .18362
misc_feature /note="match: GSS: Em:AQ151578"
19374. .19496
repeat_region /note="L2 repeat: matches 2579. .2709 of consensus"
20271. .20430
/note="MER20 repeat: matches 55. .218 of consensus"
21095. .21389
repeat_region /note="AluX repeat: matches 1. .295 of consensus"
21830. .21946
repeat_region /note="MIR repeat: matches 42. .152 of consensus"
22328. .22772
repeat_region /note="L2 repeat: matches 2242. .2710 of consensus"
22768. .22848
/note="L2 repeat: matches 2661. .2750 of consensus"
23314. .24204
/note="CpG island"
misc_feature /evidence=not_experimental
23596. .23601
misc_feature /note="Weak data."
23664. .25783
repeat_region /note="L2 repeat: matches 2588. .2741 of consensus"
25786. .26131
repeat_region /note="MER33 repeat: matches 1. .324 of consensus"
26858. .27183
repeat_region /note="AluY repeat: matches 1. .301 of consensus"
27208. .27334
/note="MIR repeat: matches 81. .221 of consensus"
27673. .27833
repeat_region /note="MIR repeat: matches 7. .175 of consensus"
complement(28112. .28516)
```

```
misc_feature /note="match: GSS: Em:AQ212338"
28535. .29088
/note="match: GSS: Em:B51862"
repeat_region 28560. .28684
/note="MIR repeat: matches 79. .210 of consensus"
28588. .28828
misc_feature /note="match: GSS: Em:B78941"
28588. .28841
misc_feature /note="match: GSS: Em:B88272"
28849. .29048
repeat_region /note="MER20 repeat: matches 2. .214 of consensus"
30148. .30232
repeat_region /note="L2 repeat: matches 2610. .2692 of consensus"
complement(31265. .31662)
misc_feature /note="match: GSS: Em:AQ625926"
/note="match: GSS: .31662)
complement(31333. .31662)
/note="match: GSS: Em:AQ625934"
complement(31491. .31857)
/note="match: GSS: Em:AQ131221"
31733. .31875
repeat_region /note="L2 repeat: matches 2039. .2197 of consensus"
33899. .33992
repeat_region /note="L2 repeat: matches 2578. .2667 of consensus"
34111. .34427
repeat_region /note="MER31A repeat: matches 4. .477 of consensus"
35235. .35300
repeat_region /note="MER5A repeat: matches 121. .189 of consensus"
35369. .35435
repeat_region
```

```
Query Match 38.0%; Score 332.8; DB 9; Length 132449;
Best Local Similarity 99.4%; Pred. No. 7.7e-77;
Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 540 gccgcctccaggaattcgaattcggccttcgaagaactgtaatgtaactgaat 599
|||||
DB 82186 GCCGCTCCGGGAATTCGAAGTTGTGGCCCTTCTCGAAGACGTATGATCCTGAAGT 82245
|||||

QY 600 ttctgaatatcgcaaacccgcaagattagcctgctgctgcgcacaaagaagaacacat 659
|||||
DB 82246 TTCGGAATATTCGCAAAACCCACAGATTAGCGTGGTCCGCAAAAGAAACACAT 82305
|||||

QY 660 agagtttaagatccagtagtgaattgaaactgttttcattgaagctgaatatata 719
|||||
DB 82306 AGAGTTTAAGATCCAGTAGAGATTGTGAACCTGTTTCATTTGAAGCTGAATATATA 82365
|||||

QY 720 cgtagtcattgttatgttggaactaaggatatactcttagcaagaagaataattccccc 779
|||||
DB 82366 CGTAGTCATGTTATGTTGAACACTAAGAGATATCTTTAGCAAGAAATATTTCCCC 82425
|||||

QY 780 ttatcccaactgctgtagagtttcttacctgcttgatgctgctgtaagaatccggga 839
|||||
DB 82426 TTATCCCACTGCTGTGAGAGTTTCTGTACCTGCTTGATGCTGTGAAGATCCGGGA 82485
|||||

QY 840 gccctgcgcaactgctgtgtgctgctgctgc 875
|||||
DB 82486 GCCTTGCCGCACTGCTGTGTGGTGGCTTGGCGCTC 82521
|||||
```

```
RESULT 9
AC044842 AC044842 172307 bp DNA HTG 25-JUN-2000
LOCUS Homo sapiens chromosome 2 clone RP11-178E20 map 2, WORKING DRAFT
DEFINITION SEQUENCE, 23 unordered pieces.
ACCESSION AC044842 GI:8705156
VERSION AC044842.2
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 172307)
Homo sapiens chromosome 2, clone RP11-178E20
```







```

Query Match      35.3%: Score 308.8; DB 6; Length 687;
Best Local Similarity 80.0%; Pred. No. 1.5e-70;
Matches 411; Conservative 0; Mismatches 52; Indels 51; Gaps 2;

OY 78 gccacatgggaatggatgaacaagatctgcgcgcctgtacatcgcgaacttcaaa 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 GCGGCATGGGAGATGGAGATGAGCCAGATCTCGCGGGCTGTACATTTGCACTTCAAA 81

OY 138 gatgcagagacgcgcgaacatctgagcaagaagaagtgacatatctgtctgcac 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 GAGGCAAGAGATGACAGAGATTGAGCGAGAGAACAGGTGACACATTTCTTCTGTGCAC 141

OY 198 gatagt-ccagacctatgttgag----- 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 GATACGTGCCAGGCCCATGTTGGAGGAGTTAAATACCTGTGTTCACAGCGGCACACA 201

OY 221 -----gacaagacattcaagaagaagatataattcaattcaagagtgccg 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 CCATCTCAAAACCTGACAGACATTTCAAGAAAGCATTAATTTCATGATGATGCCGA 261

OY 267 ctccgcggatgagagcgtgccttgtaacactgcctgcgcgggtctccagagcgctgacactg 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 CTCGAGGGGTGAGAGCTGTGTACATTGCTGCTGGGCTCCAGAGATGTGACATTTG 321

OY 327 gtgacgatacatcatgacacgctgactgtgctgtggagagatgcctgcacacacg 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 322 GTATGCGATACATCATGATGATGCTGCACGACTTTGGCTGGAGAGATCCTTGACACTGTT 381

OY 387 cgtgcgcggagatcctgtgcacaccccaacgctgtgcacagagacgtccagagatt 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 CGTGCGGGAGGTCTGTGTCACCAACCCCAACTGTGCTTCAAGGCGAGCGAGGTTT 441

OY 447 gagaagcatgagatcatcatcaatgacgtgcgtgtgaagaagaataatgagagacccct 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 GAAATACATGAAATGACCAAGATTCGGCATGCTGAGAGAGATGAGAGAACCCCT 501

OY 507 ttgcagatgcagaagaagccaaacatctcgtg 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 502 TTGCGGGATGACAGAGAACCAAAATATTCTGTG 535

RESULT 12
AC074158 198146 bp DNA HTG 02-SEP-2000
LOCUS Mus musculus clone Rp23-276B17, WORKING DRAFT SEQUENCE, 17 ordered
DEFINITION pieces
AC074158
VERSION AC074158.2 GI:9964824
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198146)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 198146)
DOE Joint Genome Institute.
Direct Submission
Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Sep 2, 2000 this sequence version replaced g1:9211268.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1842897
Center clone name: RPI-23_276B17

```

```

-----
Summary Statistics
Consensus quality: 187836 bases at least Q40
Consensus quality: 194829 bases at least Q30
Consensus quality: 196067 bases at least Q20
Estimated insert size: 20100; agarose-fp estimation
Estimated insert size: 197396; sum-of-ctdigs estimation
Quality coverage: 10.38 in Q20 bases; agarose-fp estimation
Quality coverage: 10.57 in Q20 bases; sum-of-ctdigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 52068: contig of 52068 bp in length
*
52069 52168: gap of unknown length
*
52169 53815: contig of 1647 bp in length
*
53816 53915: gap of unknown length
*
53916 55095: contig of 1180 bp in length
*
55096 55195: gap of unknown length
*
55196 59630: contig of 4435 bp in length
*
59631 59730: gap of unknown length
*
59731 74899: contig of 15169 bp in length
*
74900 74999: gap of unknown length
*
75000 83495: contig of 8496 bp in length
*
83496 83595: gap of unknown length
*
83596 92426: contig of 8831 bp in length
*
92427 92526: gap of unknown length
*
92527 93676: contig of 1150 bp in length
*
93677 93776: gap of unknown length
*
93777 108932: contig of 15156 bp in length
*
108933 109033: gap of unknown length
*
109033 113582: contig of 4550 bp in length
*
113583 113683: gap of unknown length
*
113683 124071: contig of 10389 bp in length
*
124072 124171: gap of unknown length
*
124172 145338: contig of 21067 bp in length
*
145339 145339: gap of unknown length
*
145339 164908: contig of 19570 bp in length
*
164909 165008: gap of unknown length
*
165009 174991: contig of 9983 bp in length
*
174992 175091: gap of unknown length
*
175092 181798: contig of 6707 bp in length
*
181799 181898: gap of unknown length
*
181899 197271: contig of 15373 bp in length
*
197272 197371: gap of unknown length
*
197372 198146: contig of 775 bp in length.
*
FEATURES
source
1. 198146
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="Rp23-276B17"
/clone_lib="RPI mouse BAC library 23"
BASE COUNT 53521 a 42133 c 43010 g 57881 t 1601 others
ORIGIN
Query Match      16.2%: Score 141.4; DB 2; Length 198146;
Best Local Similarity 88.0%; Pred. No. 2e-26;
Matches 154; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 293 ctgcctgcgcggatcctccagagagctgacactgtgacatcatcagaccgtcac 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167000 CAGCCTGGCTGGCGGCTCTCAGAGAGGTGACATTGTGATCCATACATCACTGTGCAC 167059

OY 353 tgaacttgctggagagatgccctgcacacgcgtgcgtgcgcggagatcgtgtgcaacc 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167060 CGACTTGGCTGGGAGAGATGCTTGCACACACTGTTGTCGGGGAGGTCTGTGCCAACCC 167119

```



```

RESULT 15
AL133540 87616 bp DNA HTG 10-JUL-2001
LOCUS AL133540/c
DEFINITION Homo sapiens chromosome 6 clone RP1-125A24 map p24.1-25.3, ***
ACCESSION AL133540
VERSION AL133540.8 GI:9884658
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 87616)
REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 87616)
JOURNAL
COMMENT
Direct Submission
Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 22, 2000 this sequence version replaced gi:9863453.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: dj125A24
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 81220 bases at least Q40
Consensus quality: 83782 bases at least Q30
Consensus quality: 85419 bases at least Q20
Insert size: 86816; sum-of-contigs
Insert size: 106856; 2.2% error; agarose-fp
Quality coverage: 2.96x in Q20 bases; sum-of-contigs Quality
coverage: 2.54x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
3452: contig of 3452 bp in length
3453 3552: gap of 100 bp
3553 23206: contig of 19654 bp in length
23207 23306: gap of 100 bp
23307 33410: contig of 10104 bp in length
33411 33510: gap of 100 bp
33511 35989: contig of 2479 bp in length
35990 36089: gap of 100 bp
36090 51429: contig of 15340 bp in length
51430 51529: gap of 100 bp
51530 56298: contig of 4769 bp in length
56299 56398: gap of 100 bp
56399 62780: contig of 6382 bp in length
62781 62880: gap of 100 bp
62881 66439: contig of 3559 bp in length
66440 66539: gap of 100 bp
66540 87616: contig of 21077 bp in length.
Location/Qualifiers
1..87616
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="p24.1-25.3"
/clone="RP1-125A24"
/clone_id="PCT-1"
1..3452
/misc_feature
/note="assembly_fragment:00717"

```

```

clone_end:r7
vector_side:left"
3553..23206
/note="assembly_fragment:00046"
23307..33410
/note="assembly_fragment:00236.0"
33511..35989
/note="assembly_fragment:00354"
36090..51429
/note="assembly_fragment:00398"
51530..56298
/note="assembly_fragment:00410"
56399..62780
/note="assembly_fragment:00571"
62881..66439
/note="assembly_fragment:00783"
66540..87616
/note="assembly_fragment:00752
clone_end:sp6
vector_side:right"
BASE COUNT 25196 a 18263 c 19295 g 24060 t 802 others
ORIGIN
Query Match 12.0%; Score 105.2; DB 2; Length 87616;
Best Local Similarity 93.2%; Pred No. 6.7e-17;
Matches 110; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 cccgcgcgtccctccctgaacatgcatagtcgctgagacacacgagccggagc 60
DB 33277 cccgcgcgtccctccctgaacatgcatagtcgctgagacacacgagccggagc 33218
QY 61 gctagcgttcgcttcagccaccatgagggatggatgaacagatcctgcggcct 118
DB 33217 gctagcgttcgcttcagccaccatgagggatggatgaacagatcctgcggcct 33160

```

Search completed: February 13, 2002, 02:10:38  
Job time: 5251 sec

Thu Feb 14 16:00:22 2002

us-09-544-525-1.rge

Page 13





DR P-PSDB; AAB23298.

XX Novel dual-specificity mitogen activated protein kinase phosphatase  
PT polypeptide useful in screening assays for identifying agents that  
PT modulate activity of the protein which are useful for treating cancer  
PS and autoimmune diseases

XX Claim 7; Fig 1; 60pp; English.

XX The invention relates to a human dual-specificity mitogen-activated  
CC protein (MAP) kinase phosphatase, DSP-3, and to nucleic acids encoding  
CC it. The invention also relates to variants of DSP-3 which retain  
CC activity, expression vectors and host cells comprising DSP-3-encoding  
CC DNA, the recombinant production of DSP-3, an anti-DSP-3 antibody, and  
CC a DSP-3 substrate-trapping mutant protein that has a reduced ability to  
CC dephosphorylate a substrate relative to the wild-type DSP-3. The  
CC invention additionally encompasses use of a DSP-3 modulator to modulate  
CC a proliferative response, cell differentiation or cell survival. The  
CC DSP-3 protein is useful for screening an agent that binds to DSP-3 and/or  
CC modulates DSP-3 activity, and is also useful for raising antibodies. DNA  
CC encoding DSP-3 or a reporter protein is also useful for screening an  
CC agent that modulates DSP-3 activity. The identified agents that  
CC modulate DSP-3 activity are useful for treating Duchenne muscular  
CC dystrophy, cancer, graft-versus-host disease, autoimmune diseases,  
CC allergies, metabolic diseases, abnormal cell growth, abnormal cell  
CC proliferation and cell cycle abnormalities. DSP-3-specific antibodies  
CC and DSP-3 antisense probes are useful for detecting DSP-3 expression in  
CC a sample. The present sequence represents cDNA encoding human DSP-3.

XX Sequence 875 BP; 218 A; 217 C; 231 G; 209 T; 0 other;

Query Match 99.8%; Score 873.4; DB 21; Length 875;

Best Local Similarity 99.9%; Pred. No. 4.8e-255;

Matches 874; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 cccgcgcgtccctccctcgtgaacatgcatagtcgcctcgcacacacgcggggc 60  
DB 1 cccgcgcgtccctccctcgtgaacatgcatagtcgcctcgcacacacgcggggc 60  
OY 61 gctagcgttcgctccacgacacatgagggaatgagatgaacaagatcctgcgcgcct 120  
DB 61 gctagcgttcgctccacgacacatgagggaatgagatgaacaagatcctgcgcgcct 120  
OY 61 gctagcgttcgctccacgacacatgagggaatgagatgaacaagatcctgcgcgcct 120  
DB 61 gctagcgttcgctccacgacacatgagggaatgagatgaacaagatcctgcgcgcct 120  
OY 121 acatcggaacttcaaaatgctcagagacgcgcgaacaattgacagaagaacagtggaac 180  
DB 121 acatcggaacttcaaaatgctcagagacgcgcgaacaattgacagaagaacagtggaac 180  
OY 181 atattctgtctgtccacagatagtcacagcctatgtgtgagagacagaacattccaagaa 240  
DB 181 atattctgtctgtccacagatagtcacagcctatgtgtgagagacagaacattccaagaa 240  
OY 241 atattctgtctgtccacagatagtcacagcctatgtgtgagagacagaacattccaagaa 240  
DB 241 atattctgtctgtccacagatagtcacagcctatgtgtgagagacagaacattccaagaa 240  
OY 241 gtattaaattcaatcaagagtcgcgcgtccgcgtgagagcctgtacactgcctcg 300  
DB 241 gtattaaattcaatcaagagtcgcgcgtccgcgtgagagcctgtacactgcctcg 300  
OY 301 ccgggggtccacagagcgtgacacgtgcatcagatcatatagccctcctgcttg 360  
DB 301 ccgggggtccacagagcgtgacacgtgcatcagatcatatagccctcctgcttg 360  
OY 361 gctggagagatgcccctgcacacgcgtgcgtgcgcgagagatcctgttccaaacccacgtcg 420  
DB 361 gctggagagatgcccctgcacacgcgtgcgtgcgcgagagatcctgttccaaacccacgtcg 420  
OY 421 gcttcacagacagctccacagagtttgaagaacatgagtcacatcagatcgcagtcg 480  
DB 421 gcttcacagacagctccacagagtttgaagaacatgagtcacatcagatcgcagtcg 480  
OY 481 tgaaggaagaatataggaagagccttgcagagatgcagaagaacgcaaaacattctcg 540  
DB 481 tgaaggaagaatataggaagagccttgcagagatgcagaagaacgcaaaacattctcg 540  
OY 541 ccgctccaggaattctgagttcgtgaccttctcagaagaagcctgtaagtacccctgaagt 600

DB 541 cccgtccaggaattctgagttcgtgaccttctcagaagaagcctgtaagtacccctgaagt 600  
OY 601 tctgaataatctgcaaacccgcagagtttagctgtgtcgtcccaaaaagaacacata 660  
DB 601 tctgaataatctgcaaacccgcagagtttagctgtgtcgtcccaaaaagaacacata 660  
OY 661 gagtttaagtaaccagtagtattgttaactgttttcatcttgaagtcgatatatac 720  
DB 661 gagtttaagtaaccagtagtattgttaactgttttcatcttgaagtcgatatatac 720  
OY 721 gtagtcaatcttattgtgagacataagatatctttagcaagaagaataattccct 780  
DB 721 gtagtcaatcttattgtgagacataagatatctttagcaagaagaataattccct 780  
OY 781 tatcccaactgtgtgagagttctgtacactgccttgatgacgtcgtgaaagaccggag 840  
DB 781 tatcccaactgtgtgagagttctgtacactgccttgatgacgtcgtgaaagaccggag 840  
OY 841 ccttgcgcacatgccttctgtggtgtggtgcctc 875  
DB 841 ccttgcgcacatgccttctgtggtgtggtgcctc 875

RESULT 2

ID AAF32191 standard; CDNA; 926 BP.

XX AAF32191;

XX AAF32191;

XX 12-APR-2001 (first entry)

XX Human dual-specificity phosphatase DSP-3 coding sequence.

XX Human: DSP-3; dual-specificity phosphatase; cell proliferation;

XX cell signalling; cancer; graft-versus-host disease; autoimmune disease;

XX allergy; metabolic disease; Duchenne muscular dystrophy; ss.

XX Homo sapiens.

XX WO200102581-A1.

XX 11-JAN-2001.

XX 20-APR-2000; 2000WO-US10868.

XX 02-JUL-1999; 99US-0142338.

XX 07-APR-2000; 2000WO-US09185.

XX (CEPT-) CEPTYR INC.

XX Lucite RM, Wei B;

XX WPI; 2001-138148/14.

XX P-PSDB; AAB67167.

XX New dual-specificity phosphatase-3 polypeptide and its variants useful

XX for treating disorders associated with DSP-3 activity, defects in cell

XX proliferation, differentiation or survival, e.g. Duchenne muscular

XX dystrophy, cancer

XX Claim 7; Fig 1; 70pp; English.

XX The present invention provides the protein and coding sequences of the

XX human dual-specificity phosphatase DSP-3. The DSP-3 protein is involved

XX in cell signalling and the sequences can be used in the treatment of

XX cancer, metabolic and autoimmune diseases, allergies, graft-versus-host

XX disease, abnormal cell proliferation and Duchenne muscular dystrophy.

XX Sequence 926 BP; 233 A; 230 C; 242 G; 221 T; 0 other;

XX Query Match 91.7%; Score 802.4; DB 22; Length 926;





```

QY 221 -----gacaagacattccaagaagatataat 249
DB 241 tcccaagcagcgaattccacatctcaaaacctgacaagacattccaagaagatataat 300
QY 250 tcaattacagagtgccgcgtccgcggtgagagcgtgcttgtaacctgtccctggccggggtct 309
DB 301 tcaattacagagtgccgcgtccgcggtgagagcgtgcttgtaacctgtccctggccggggtct 360
QY 310 ccagagcgtgacacgtgcatcgatcacatcatgacccgtcactgtgcttggctggagag 369
DB 361 ccagagcgtgacacgtgcatcgatcacatcatgacccgtcactgtgcttggctggagag 420
QY 370 atgacctgcaacacgtgcatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgat 429
DB 421 atgacctgcaacacgtgcatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgat 480
QY 430 gacagctcagagagtgtagaagcagcagcagcagcagcagcagcagcagcagcagcagcag 489
DB 481 gacagctcagagagtgtagaagcagcagcagcagcagcagcagcagcagcagcagcagcag 540
QY 490 aatatgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 549
DB 541 aatatgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 600
QY 550 gaattctgaagctcggcccttctcaagaagactgtaactgtaactgtaactgtaactgtaact 609
DB 601 gaattctgaagctcggcccttctcaagaagactgtaactgtaactgtaactgtaactgtaact 660
QY 610 ttgcaaacccgcagagtttagagtgctgctgcaaaaaaagaagcaacatagagtttaag 669
DB 661 ttgcaaacccgcagagtttagagtgctgctgcaaaaaaagaagcaacatagagtttaag 720
QY 670 tatccagtagtgattgtaactgtgtttcatttgaagctgtaataatactacgtacgtacgt 729
DB 721 tatccagtagtgattgtaactgtgtgtttcatttgaagctgtaataatactacgtacgtacgt 780
QY 730 ttatgttgagaactaagaatattctttagcaagaagaataatttcccttatcccccac 789
DB 781 ttatgttgagaactaagaatattctttagcaagaagaataatttcccttatcccccac 840
QY 790 tgcgtgagaggtttctgtacccctgctgtaagcgtgtaagatcccggaagccttgccgcg 849
DB 841 tgcgtgagaggtttctgtacccctgctgtaagcgtgtaagatcccggaagccttgccgcg 900
QY 850 actgccttggtggtgctggtgcgcgc 875
DB 901 actgccttggtggtgctggtgcgcgc 926

```

RESULT 4  
AAF63568  
ID AAF63568 standard; cDNA; 1379 BP.

AC AAF63568;

DT 11-MAY-2001 (first entry)

DE Human phosphatase AA374753.h coding sequence.

XX Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;  
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;  
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;  
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;  
KW schizophrenia; hamartoma; ss.

XX Homo sapiens.

XX WO200112819-A2.

XX 22-FEB-2001.

XX 11-AUG-2000; 2000WO-US22158.

```

XX 13-AUG-1999; 99US-0149005.
PR (SUGEN-) SUGEN INC.
XX
XX Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;
PI WPI: 2001-211226/21.
XX P-PSDB; AAB73216.
DR
DR New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
XX Claim 1; Fig 4; 138pp; English.
XX
XX The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase coding sequence.
CC Phosphatases are enzymes that catalyse the dephosphorylation of proteins
CC modified by phosphorylation of serine, threonine or tyrosine residues.
CC The phosphatases are useful for treating a variety of diseases: for
CC example cancer e.g. breast, urogenital, prostate, head, neck, lung
CC cancers, synovial sarcomas, renal cell carcinoma, non-small cell lung
CC cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach
CC cancer, glioblastoma, colorectal cancer and thyroid cancer,
CC pathophysiological hypoxia, cardiac dysfunction and/or vascular
CC disorders, myopathies, congenital muscle disorders, Papillon-Lefevre
CC syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome,
CC Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and
CC hamartomas.
XX
XX Sequence 1379 BP; 298 A; 374 C; 386 G; 320 T; 1 other;
SQ

```

Query Match 91.3%; Score 799.2; DB 22; Length 1379;  
Best Local Similarity 94.2%; Pred. No. 2,1e-232;  
Matches 872; Conservative 0; Mismatches 3; Indels 51; Gaps 2;

```

QY 1 cccgcgcgtctctcccttgaatgcatgagcgtgctgcaacacagcgcgggac 60
DB 257 cccgcgcgtctctcccttgaatgcatgagcgtgctgcaacacagcgcgggac 316
QY 61 gctagcgttgccttcaagccacatgaggaatgagatgaagaatctgcccgcgtgt 120
DB 317 gctagcgttgccttcaagccacatgaggaatgagatgaagaatctgcccgcgtgt 376
QY 121 acatcggcaacttcaagaatgctcagagagcgcggaacattgagcaagaacagtgacac 180
DB 377 acatcggcaacttcaagaatgctcagagagcgcggaacattgagcaagaacagtgacac 436
QY 181 atattcgtctgtccagatagt-ccagggcctattgtgag----- 220
DB 437 atattcgtctgtccagatagtgtccagggcctattgtgagggagttaatcctgtgca 496
QY 221 -----gacaagacattccaagaagatataat 249
DB 497 tcccaagcagcgaattccacatctcaaaacctgacaagaacattccaagaagatataat 556
QY 250 tcaattacagagtgccgcgtccgcggtgagagcgtgcttgtaacctgtccctggccggggtct 309
DB 557 tcaattacagagtgccgcgtccgcggtgagagcgtgcttgtaacctgtccctggccggggtct 616
QY 310 ccagagcgtgacacgtgcatcgatcacatcatgacccgtcactgtgcttggctggagag 369
DB 617 ccagagcgtgacacgtgcatcgatcacatcatgacccgtcactgtgcttggctggagag 676
QY 370 atgacctgcaacacgtgcatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgat 429
DB 677 atgacctgcaacacgtgcatcgatcgatcgatcgatcgatcgatcgatcgatcgatcgat 736
QY 430 gacagctcagagagtgtagaagcagcagcagcagcagcagcagcagcagcagcagcagcag 489
DB 737 gacagctcagagagtgtagaagcagcagcagcagcagcagcagcagcagcagcagcagcag 796

```

QY 490 aataatgagagagcccttcagatgcagaagaacccaacatttcgcccctcag 549  
 |||||||  
 Db 797 aataatgagagagcccttcagatgcagaagaacccaacatttcgcccctcag 856  
 QY 550 gaattctgaagttctggccttcctcagaagctgtaattacatgaagttctgaata 609  
 |||||||  
 Db 857 gaattctgaagttctggccttcctcagaagctgtaattacatgaagttctgaata 916  
 QY 610 ttgcaaaccccgagatttagcgtgctgcacaaagaagaacacatagagtttaag 669  
 |||||||  
 Db 917 ttgcaaaccccgagatttagcgtgctgcacaaagaagaacacatagagtttaag 976  
 QY 670 tatccagtagtattgttaaaactgttttcatttgaaagctgaataataatcgtatcgt 729  
 |||||||  
 Db 977 tatccagtagtattgttaaaactgttttcatttgaaagctgaataataatcgtatcgt 1036  
 QY 730 tttaattgtgagaactaagatattctttaagaagaataatttccttccttcacac 789  
 |||||||  
 Db 1037 tttaattgtgagaactaagatattctttaagaagaataatttccttccttcacac 1096  
 QY 790 tgcctgtagaagttctctacatgcgtgtgatactgtaagaatcccgagacgttcgcgc 849  
 |||||||  
 Db 1097 tgcctgtagaagttctctacatgcgtgtgatactgtaagaatcccgagacgttcgcgc 1156  
 QY 850 actgcctgtggttgcttgagcgtc 875  
 |||||||  
 Db 1157 actgcctgtggttgcttgagcgtc 1182  
 RESULT 5  
 AAZ46164  
 ID AAZ46164 standard; cDNA: 1290 BP.  
 AC AAZ46164;  
 XX  
 DT 16-MAY-2000 (first entry)  
 XX  
 DE cDNA sequence encoding a human phosphorylation effector PHSP-27.  
 XX  
 KW Human; phosphorylation effector; PHSP; proliferative disorder;  
 KM Immune disorder; neuronal disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 238..792  
 FT /\*tag= a  
 FT /product= "phosphorylation effector"  
 XX  
 PN WO200006728-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 28-JUL-1999; 99MO-US17132.  
 XX  
 PR 28-JUL-1998; 98US-0123494.  
 PR 14-SEP-1998; 98US-0152814.  
 PR 14-OCT-1998; 98US-0173482.  
 PR 03-NOV-1998; 98US-0106889.  
 PR 19-NOV-1998; 98US-0109093.  
 PR 22-DEC-1998; 98US-0113796.  
 PR 12-JAN-1999; 98US-0173482.  
 PR 12-JAN-1999; 98US-0229005.  
 XX  
 PA (INCYTE PHARM INC.  
 XX  
 PI Hillman JL, Lai P, Tang YT, Corley NC, Guegler KJ, Baughn MR;  
 PI Patterson C, Bandman O, Au-Young J, Gorgone GA, Yue H, Azimzai Y;  
 PI Reddy R, Lu DAM, Shih LL;  
 XX  
 DR WPI: 2000-18125/16.  
 DR P-PSDB: AAY68795.

XX  
 PT New human phosphorylation effectors useful for the diagnosis, treatment  
 PT and prevention of proliferative, immune and neuronal disorders -  
 XX  
 PS Claim 9; Page 139; 142pp; English.  
 XX  
 CC AAZ46138-246168 encode human phosphorylation effectors (PHSP),  
 CC designated PHSP1-PHSP1 (the protein sequence for PHSP28 is not given  
 CC in the specification). The sequences were isolated from cDNA libraries  
 CC prepared from various human tissues. The PHSP proteins are useful for  
 CC the diagnosis, treatment and prevention of proliferative disorders,  
 CC immune disorders and neuronal disorders. The PHSP proteins form  
 CC pharmaceutical compositions which useful for treating or preventing  
 CC disorders associated with decreased PHSP expression/activity. PHSP  
 CC antagonists are useful for treating or preventing disorders associated  
 CC with increased PHSP expression/activity.  
 XX  
 SQ Sequence 1290 BP; 299 A; 341 C; 339 G; 311 T; 0 other;  
 Query Match 91.2%; Score 797.6; DB 21; Length 1290;  
 Best Local Similarity 94.1%; Pred. No. 6.3e-232;  
 Matches 871; Conservative 0; Mismatches 4; Indels 51; Gaps 2;  
 QY 1 ccccgccgctctctctccctgtaacatgcatatgctgcgcacacacgcccgggc 60  
 |||||||  
 Db 155 ccccgccgctctctctccctgtaacatgcatatgctgcgcacacacgcccgggc 214  
 QY 61 gctagcgttcgcttcagccacatgaggaatgataaagaatcctgcgcgctgt 120  
 |||||||  
 Db 215 gctagcgttcgcttcagccacatgaggaatgataaagaatcctgcgcgctgt 274  
 QY 121 acatcggaacttcaagaatgccagagacgagacaalttagacaagaatgacac 180  
 |||||||  
 Db 275 acatcggaacttcaagaatgccagagacgagacaalttagacaagaatgacac 334  
 QY 181 atattctgtctgtccacagatgt-ccaagcctatgttggag----- 220  
 |||||||  
 Db 335 atattctgtctgtccacagatgtgtccaagcctatgttggagaggttaaatcgttga 394  
 QY 221 -----gacaagaactttaaagaagttaaat 249  
 |||||||  
 Db 395 tcccaagacgagattccacatcacaatcaaaactgacaagaacattcaagaagtattaa 454  
 QY 250 tcaattcaagatgcccgcgtccgcgtgtgagagctgctgtgaactgctgcgcgggtct 309  
 |||||||  
 Db 455 tcaattcaagatgcccgcgtccgcgtgtgagagctgctgtgaactgctgcgcgggtct 514  
 QY 310 ccaagagcgtgacacactgtgtgacatcacatcacatgacccgtgacttggctggag 369  
 |||||||  
 Db 515 ccaagagcgtgacacactgtgtgacatcacatcacatgacccgtgacttggctggag 574  
 QY 370 atgacctgacacacgctgcgtgcgcggagatcctgtgcaaccccaacgctggctcaga 429  
 |||||||  
 Db 575 atgacctgacacacgctgcgtgcgcggagatcctgtgcaaccccaacgctggctcaga 634  
 QY 430 gacagctcccaagattttagaagaatgagtgccatcaatcgtgagagtgcttgaagag 489  
 |||||||  
 Db 635 gacagctcccaagattttagaagaatgagtgccatcaatcgtgagagtgcttgaagag 694  
 QY 490 aatatgagagagcccttcagagatgacagagaagccaacaaacttgcgcgtccag 549  
 |||||||  
 Db 695 aatatgagagagcccttcagagatgacagagaagccaacaaacttgcgcgtccag 754  
 QY 550 gaattctgaagttctggccttcctcagaagctgtaattacatgaagttctgaata 609  
 |||||||  
 Db 755 gaattctgaagttctggccttcctcagaagctgtaattacatgaagttctgaata 814  
 QY 610 ttgcaaaccccgagatttagcgtgctgcacaaagaagaacacatagagtttaag 669  
 |||||||  
 Db 815 ttgcaaaccccgagatttagcgtgctgcacaaagaagaacacatagagtttaag 874  
 QY 670 tatccagtagtattgttaaaactgttttcatttgaaagctgaataataatcgtatcgt 729







CC protein has functional activities comprising dual-specificity phosphatases  
CC activity and TNK pathway upregulation activity. The IKAP polypeptide can

KW   vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective

KW   vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective

KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiast;



immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 hypotensive; dermatological; immunosuppressive; antineoplastic;  
 antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;  
 antianemic; gene therapy; cancer; proliferative disorder; hypotension;  
 neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 cholesterol ester storage; systemic lupus erythematosus; infection;  
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 bone damage; cartilage damage; antineoplastic disease; coagulation;  
 thrombosis; contraceptive; ss.  
 Homo sapiens.  
 WO200058473-A2.  
 05-OCT-2000.  
 31-MAR-2000; 2000WO-US08621.  
 31-MAR-1999; 99US-0127607.  
 02-APR-1999; 99US-0127636.  
 05-APR-1999; 99US-0127728.  
 30-MAR-2000; 2000US-0540763.  
 (CURA-) CURAGEN CORP.  
 Shinkens RA, Leach M;  
 WPI: 2000-602362/57.  
 P-PSDB; AAB41570.  
 Novel nucleic acids and peptides derived from open reading frame X,  
 useful for treating e.g. cancers, proliferative disorders,  
 neurodegenerative disorders and cardiovascular disease -  
 Claim 5; Page 1910-1911; 5507PP; English.  
 AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 which represent the human ORF open reading frames 1 to 3161. The ORF  
 sequences have activities such as: cytosolic; hepatotropic; vulnary;  
 antiproliferative; antiparkinsonian; nocrotropic; neuroprotective;  
 osteoplastic; anticonvulsant; antirheumatic; immunosuppressive;  
 immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 antidiabetic; hypotensive; dermatological; immunosuppressive;  
 antineoplastic; antibacterial; antiviral; antifungal; antirheumatic;  
 antihypertensive; antianemic. The sequences can be used for determining  
 the presence of or predisposition to, or preventing or treating  
 pathological conditions associated with an ORF-associated disorder. The  
 nucleic acids can be used to express ORF proteins in gene therapy  
 vectors. The proteins and nucleic acids may be used to treat cancers,  
 proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 graft vs host disease, cardiovascular disease, diabetes mellitus,  
 hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 nocturnal haemoglobinuria, antineoplastic disease, to enhance  
 coagulation; to inhibit thrombosis; and as a contraceptive.  
 Sequence 289 BP; 71 A; 71 C; 84 G; 62 T; 1 other;  
 Query Match 28 3%; Score 247.6; DB 21; Length 289;  
 Best Local Similarity 95.8%; Fred. No. 3e-65;  
 Matches 276; Conservative 0; Mismatches 9; Indels 3; Gaps 2;  
 82 ccatgggaatggatgacaagaatccgcccgcctgtacatggcgaattcaagatg 141  
 |||||||  
 Db 2 ccatgggaatggatgacaagaatccgcccgcctgtacatggcgaattcaagatg 61  
 |||||||  
 142 ccagagacgcggaacatgtgacgaagaagtg-acacatattctgtccacat 200  
 |||||||

Db 62 ccagagacgcggaacatgtgacgaagaaggggaaccctttctgttcccgat 121  
 QY 201 ag- tccagcctatgttggaggaacagacatttcaaaagaagtataatcattcaag 258  
 |||||||  
 Pb 122 ggggtccagagcctatgttggaggaacagacatttcaaaagaagtataatcattcaag 181  
 QY 259 agtgcgcgtccgcggtgagagctgctgtacacgtgcgcggggtcccaagagcg 318  
 |||||||  
 Db 182 agtgcgcgtccgcggtgagagctgctgtacacgtgcgcggggtcccaagagcg 241  
 QY 319 tgacactgtgatacgaatacgaatgacgcgtgactgtgactgtgactgtgactgtg 366  
 |||||||  
 Db 242 tgacactgtgatacgaatacgaatgacgcgtgactgtgactgtgactgtgactgtg 289  
 RESULT 12  
 AAD09498  
 ID AAD09498 standard; DNA; 1326 BP.  
 AC AAD09498;  
 DT 10-SEP-2001 (first entry)  
 XX Human SGP008 phosphatase polypeptide encoding DNA.  
 DE  
 XX Human: SGP008 phosphatase polypeptide: phosphatase-related disease;  
 KW immune-related disorder: ocular disease; organ transplant rejection;  
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiant;  
 KW parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
 KW attention disorder; cognition disorder; psychotic disorder; cytosolic;  
 KW neurological disorder; virinicide; nocrotropic; cerebroprotective; therapy;  
 KW neuroprotective; antibacterial; vulnary; tranquilliser; antistimulant;  
 KW hypotensive; immunosuppressive; antiproliferative; analgesic; hypertensive;  
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;  
 KW MKP; migraine; chromosome 20q11.2; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..971  
 FT CDS  
 FT /\*tag= a  
 FT /product= "Human SGP008 phosphatase polypeptide"  
 FT /transl\_except= (pos:967..971, aa:LSKKGR)  
 FT /note= "CDS does not include stop codon"  
 FT /partial  
 XX  
 PN WO200146394-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 21-DEC-2000; 2000WO-US34736.  
 XX  
 PR 21-DEC-1999; 99US-0173255.  
 XX  
 PR 28-DEC-1999; 99US-0175766.  
 XX  
 PR 25-JAN-2000; 2000US-0178078.  
 XX  
 PR 31-JAN-2000; 2000US-0179301.  
 XX  
 PA (SUGEN-) SUGEN INC.  
 XX  
 PI Plowman GP, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;  
 XX Flanagan P;  
 XX WPI: 2001-418058/44.  
 DR P-PSDB; AAE04840.  
 XX  
 XX Novel phosphatase polypeptide useful for treating cancers,  
 PT immune-related diseases and disorders, cardiovascular disease, brain or  
 PT neuronal-associated diseases and metabolic disorders -  
 XX Claim 29; Fig 1; 186pp; English.  
 XX





KW Rheumatoid arthritis; microbial infection; trauma; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT 172..837  
 FT CDS /tag= a  
 FT /product= "regulator of intracellular phosphorylation"  
 PN W0200055332-A2.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 17-MAR-2000; 2000WO-US07277.  
 XX  
 PR 18-MAR-1999; 99US-0125593.  
 PR 20-MAY-1999; 99US-0135049.  
 PR 09-JUL-1999; 99US-0143188.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
 PI Lu DAM, Au-Young J;  
 DR WPI: 2000-602121/57.  
 DR P-PSDB: AAB18667.  
 XX  
 PT Novel human intracellular phosphorylation regulator polypeptides and  
 PT polynucleotides for diagnosis, prevention and treatment of  
 PT neurological, cell proliferative and autoimmune/inflammatory disorders  
 PT -  
 PT  
 PS  
 PS Claim 4; Page 94; 96pp; English.  
 CC The present sequence encodes a human regulator of intracellular  
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and  
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
 CC are useful for treating a disease or condition associated with  
 CC decreased or increased expression of functional HRIP. Diseases treated  
 CC or diagnosed include neurological disorders such as stroke, Parkinson's  
 CC disease, demyelinating diseases, bacterial and viral meningitis and  
 CC other developmental disorders of the central nervous system,  
 CC neuromuscular disorders, masthenia gravis, cell proliferative disorders  
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,  
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
 CC inflammatory disorder such as Addison's disease, acquired  
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
 CC rheumatoid arthritis, microbial infection and trauma.  
 CC  
 XX  
 XX Sequence 1212 BP; 284 A; 345 C; 373 G; 210 T; 0 other:  
 SO

Query Match 6.7%; Score 58.6; DB 21; Length 1212;  
 Best Local Similarity 57.6%; Pred. No. 1.8e-07;  
 Matches 125; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

OY 284 cctgttaacactgctgcccgggtctccagagcgttgacaactgtgatacatatcat 343  
 ||| || ||||| || || || || ||||| || ||||| ||  
 DB 603 cctgttcaactgcgtacatggcgcagcgcgcacccctgtccctgactcatg 662  
 ||| || ||||| || || || || ||||| || ||||| ||  
 OY 344 gacgcgtgactgtgctgtgagagatgacctgacacgcgtgctgcgcgagatctctg 403  
 ||| || ||||| || || || || ||||| || ||||| ||  
 DB 663 gatccacaagacatgacacctgtgtgagacgcatccacagaatg---gccaaagaccgctg 719  
 ||| || ||||| || || || || ||||| || ||||| ||  
 OY 404 tgcacaccaccaacgtggcgtccacagacagctccagagatttgagaagcatgaggtcca 463  
 ||| || ||||| || ||||| || ||||| || ||||| ||  
 DB 720 cgtctcccgacacggggtcttttgaagcagctccggagagctggacaagacagctgtgca 779  
 ||| || ||||| || || || || ||||| || ||||| ||  
 OY 464 tcaatatacgcagtgctgaaagaaatataatgagag 500  
 ||| || ||||| || || || || ||||| || ||||| ||  
 DB 780 gcagagcgacggtccccagcgccagacgagctgagag 816

RESULT 15  
 AAD09495  
 ID AAD09495 standard; DNA; 1262 BP.  
 XX  
 XX  
 AC AAD09495;  
 XX  
 DT 10-SEP-2001 (first entry)  
 XX  
 DE Human SGP003 phosphatase polypeptide encoding DNA.  
 XX  
 KW Human; SGP003 phosphatase polypeptide; phosphatase-related disease;  
 KW immune-related disorder; ocular disease; organ transplant rejection;  
 KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;  
 KW metabolic disorder; haematopoietic cancer; mood disorder; cardiac;  
 KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;  
 KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;  
 KW attention disorder; cognition disorder; psychotic disorder; cytostatic;  
 KW neurological disorder; vincristine; neurotropic; cerebroprotective; therapy;  
 KW neuroprotective; antibacterial; vulnerrary; tranquiliser; antispasmodic;  
 KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;  
 KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;  
 KW MKP; migraine; chromosome CHR10; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 240..902  
 FT CDS /tag= a  
 FT /product= "Human SGP003 phosphatase polypeptide"  
 FT /trans\_except= (pos:492..494, aa:Leu)  
 FT  
 FT  
 FT  
 FT  
 PN W0200146394-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 21-DEC-2000; 2000WO-US34736.  
 XX  
 PR 21-DEC-1999; 99US-0173255.  
 PR 28-DEC-1999; 99US-0175766.  
 PR 25-JAN-2000; 2000US-0178078.  
 PR 31-JAN-2000; 2000US-0179301.  
 XX  
 PA (SUGEN) SUGEN INC.  
 XX  
 PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;  
 PI Flanagan P;  
 PI  
 DR WPI: 2001-418058/44.  
 DR P-PSDB: AAB04837.  
 XX  
 XX Novel phosphatase polypeptide useful for treating cancers,  
 PT immune-related diseases and disorders, cardiovascular disease, brain or  
 PT neuronal-associated diseases and metabolic disorders  
 PT  
 PS  
 PS Claim 29; Fig 1; 186pp; English.  
 XX

The present invention relates to phosphatase polypeptides, nucleotide sequences encoding them, as well as various products and methods useful for the diagnosis and treatment of various phosphatase-related diseases and conditions. Substance that modulates the activity of phosphatase polypeptide is used to treat immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases and metabolic disorders, including cancers of tissues, cancers of haematopoietic origin, diseases of central and peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders, dyskinesias and organ transplant rejection. The present sequence is a DNA encoding human SGP003 phosphatase polypeptide. This sequence is classified as dual specificity phosphatase (DSP) and MAP

CC Kinase phosphatase (MKP). SGP003 gene maps to chromosomal position  
CC CHR10.

XX  
SQ Sequence 1262 BP; 287 A; 356 C; 392 G; 227 T; 0 other;

Query Match 6.7%; Score 58.6; DB 22; Length 1262;  
Best Local Similarity 57.6%; Pred. No. 1.8e-07;  
Matches 125; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

```
QY 284 cctgtacactgctgcccggggtctccaggagcgtgacactggtgatgcatacatcat 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 668 cctgttccactgctcatatgggcccggcggcgtcagccacctggtccttgctactgat 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 gaccgtacactgacttggcttggaggatgacccctgcacacgctgcgtgccgggagatccctg 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 728 gatccacaaggacatlgaccctcgtgtggaagcgcacatccagcaagtg--gccaaagaaccgctg 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 tgcacaaccacaagctgggcttccagagacagctccagagagtttgagaagcatgagtcaca 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 cgtctctccggaaaccgggcttlttgaagcagctccgggagcttgacaagcagctgltgca 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 tcagtatcggcagctggtgctgaaggaagaatatgagagag 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 gcagagggcgacggtcccaagcggcaggaagcgtcgaaggaag 881
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 13, 2002, 02:11:26  
Job time: 3113 sec





BASE COUNT 205 a 160 c 182 g 170 t  
 ORIGIN

Query Match 67.3%; Score 588.6; DB 10; Length 717;  
 Best Local Similarity 92.3%; Pred. No. 6.3e-156;  
 Matches 662; Conservative 0; Mismatches 4; Indels 51; Gaps 2;

100 acagaatctgcccgcgcgtatcatggaactcaagaatgccagaagcgggaacat 159  
 |||||||  
 Db 1 ACAGATCTCTGCCCGCCTTACATGGCACTTCAAGATGCCAGAGACGCCGAACAT 60  
 160 tgaagcaagaacaagtgacacatatctgtctgcacagatg-ccaagccatattg 218  
 |||||||  
 Db 61 TGAGCAAGAACAAGTGACACATATCTGCTGTCCACGATAGTGCAGGCCATATG 120  
 219 ag-----gacaagac 228  
 ||  
 Db 121 AGGAGATTAAATACCTGTGATCCAGACGCGATTCCACATCTCAAAACCTGACAAGAC 180  
 229 attcaagaagaattataatcatcagagtgccggtccggtgagagctgcttg 288  
 |||||||  
 Db 181 ATTTCAAAGAAATATTAAATTCATTCAGAGTCCGCGCTCCGCGTGAGAGCTGCTTG 240  
 289 taactgcttgccggggtctccaaggagcgtgacactggtgacatcatcatgacg 348  
 |||||||  
 Db 241 TACACTGCTGGCGGGGCTCTCCAGAGAGCGTGACACTGATGATCATCATATACCG 300  
 349 taactgacttgctgaggagagatgcccgcacacccgtgctgcccggagatccctgtgca 408  
 |||||||  
 Db 301 TCAGTGACTTTGGCTGGGAGATGCCCTGCACACCGCTGCTGGGAGATCCTGTGCCA 360  
 409 accccaagctgggcttcagaagacagctccagaagtttgaagaagcagtggtccatcaat 468  
 |||||||  
 Db 361 ACCCCACAGCTGGGCTTCCAGAGACAGCTCCAGAGATTGAAGAGCATGAGTCCATCACT 420  
 469 atcgcagctgctggaagagaataatgagagagagcccttgcagagtgcaagaagaagcca 528  
 |||||||  
 Db 421 ATCGGAGTGGCTGAAGAGAAATATGAGAGAGCGCTTTCAGAGATGAGAGAAAGCCA 480  
 529 aaaaacatctgcccgcgtccagaagatctgaaagttctgggcttctccagaagactgtaat 588  
 |||||||  
 Db 481 AAAACATTTCTGGCCGCTCCAGAGATTTCTGAAGTTTGGCGCTTTCAGAAAGCTGTAT 540  
 589 gfacctgaagttctgaaatattgaaacccgcagaagtttaagccggtgctgcccacaaag 648  
 |||||||  
 Db 541 GTACCTGAAGTTTCTGAATATTTGCAAAACCCACAGAGTTTAAAGCTGTGCTCCCAAAAG 600  
 649 aaaagcaacatagatttaagatcagtagagattgtaactgatttcatctgaag 708  
 |||||||  
 Db 601 AAAAGCAACATAGAGTTTAAATTCACAGAGTGAATTTGAAACCTGTTTATTTTGAAG 660  
 709 ctgaatatatacgtatgcatgttattgttgaagactaaagatatcttaagcaagag 765  
 |||||||  
 Db 661 CTGAATATATACGTAGTATGTTTGTGAGAACTAAAGATATCTTATAGCAAGAG 717

RESULT 2  
 LOCUS BE868556 915 bp mRNA EST 20-OCT-2000  
 DEFINITION 601444589P1 NIH\_MGC\_65 Homo sapiens cDNA clone IMAGE:3848374 5',  
 mRNA sequence.  
 ACCESSION BE868556  
 VERSION BE868556  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 915)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE

JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgraphs-remail.nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM9564 row: b column: 23  
 High quality sequence stop: 693.  
 Location/Qualifiers

## FEATURES

source

1..915

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3848374"

/clone\_lib="NIH\_MGC\_65"

/tissue\_type="adenocarcinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: Colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.8 kb. Library constructed by Life  
 Technologies."

BASE COUNT 248 a 221 c 242 g 204 t  
 ORIGIN

Query Match 63.9%; Score 559; DB 11; Length 915;  
 Best Local Similarity 86.3%; Pred. No. 1.6e-147;  
 Matches 698; Conservative 0; Mismatches 55; Indels 56; Gaps 5;

86 gggaatgagatgaacaatctctgcccgcgtgtacatcgcaacttaagaagcag 145  
 |||||||  
 Db 8 GGGGAATGGGATGAACAAAGATCTCGCCGCTGTACATCGGCAACTCAAGATCCAG 67  
 146 aagacgcgaacaattgaagaagaagaagtgacatatctgtctgtccacagatg-c 204  
 |||||||  
 Db 68 AGACGGGACAACTTGTAGCAAGAAGGTGACACATATTTCTGTCTGTCCAGATATGTC 127  
 205 caggccatgttgag----- 220  
 |||||||  
 Db 128 CAGGCTTATGTTGGAGGGAGTTAAATACCTGTGATCCCATCAGCGGATTCACATCTCA 187  
 221 -----gaaagacattccaagaagaatataatcatcagagtgccggtccgagg 274  
 |||||||  
 Db 188 AAACCTGACAAAGACATTTTCAAGAAAGTAATTAATTCATTCACGAGTCCGCGCTCCGGG 247  
 275 tgaagctgacctgtgacactgctgcccggggtctccagaagcgtgacactggtgacgc 334  
 |||||||  
 Db 248 TGAGAGCTGCTTGTACACTGCTGCGCGGCTTCCAGAGACGTGACACTGATGATGCG 307  
 335 atacatcatgacgtcactgacttggctggtggaagatgccctgcaacacgctgctgcccg 394  
 |||||||  
 Db 308 ATACATCATGACCGTACACTGACTTGGCTGGGAGGATGCCCTGCACACCGTGGCTGTG 367  
 395 ggaatcctgtgccaaccccaacgctggtctccagaagacgtccagaagattgaaga 454  
 |||||||  
 Db 368 GAGATCTGTGCAACCCCAACGTTGGCTTCCAGAGACAGTCCAGGAGTTTGAAGAACCA 427  
 428 tgaagtcacatgactgagcagtgctgaagaagaatatggaagagcccttgcaga 514  
 |||||||  
 Db 428 TGAGTCCATCATCTATCGGCACTGCTTAAGAGAAATATGAGAGAGCCCTTGCACAGA 487  
 515 tgcagaagaagccaacaaatcttgcgcgtccagaagattcgaagttctggtgcttct 574  
 |||||||  
 Db 488 TGCAAGAAGAGCAAAACATTTCTGCGCTCCGGAATTCGAAGTTCTGGGCTTTCT 547  
 575 cagaagactgtatgactcctgaagtttclgaaatatgcaaacccgcagaagattgagctg 634  
 |||||||  
 Db 548 CAGAAGACTGTAAATGTACTCAAGTTTCTGAATATGCAAGACACACAGAGTTTAAAGCTG 607  
 635 gtgctgcacaaaagaagaacacatagatgattgaagttacacagtgatgtaactgt 694

Db	121	GACAGCTCCAGAGATTGTGAGAACGATGAGCTTCATCAGTATCGGCACAGTGCCTGAAGGAAG	180
Qy	490	aatatlgagagaagcccttltgagagatlgcaagaagaagccaaaaattctgacgctccag	549
Db	181	AATATGAGAGAGAGCCCTTTGAGAGATGAGAGAGAGCCAAACATCTCTGCCCCCTCCAG	240
Qy	550	gaattctgaagttcttgagcccttctcagaagaacgttaatgtaacctgaagttcttgaata	609
Db	241	GAATTTCTGAAGTTCTGGGCCCTTTCTCTCAGAACACTGTAAATGTAAGTCTGAAGTTCTGAAATA	300
Qy	610	ttggaacaccgcagaagtttaagctggtgctgctgcacaaaagaagaacaatagatlttaag	669
Db	301	TTGCAAAACCCACAGAGTTTAGGTGTGCTGCTCCCAAAAAGAAAGCAATAGAGTTTAAG	360
Qy	670	tatccagtagtgatltgttaaacctglttltcaatttgaagctgaatatataacgtagcatg	729
Db	361	TATCCAGTAGAGATTTGTAAACCTGTCTTTTCAATTTGAAGCTGAATAATATAGTACTCATG	420
Qy	730	tttatgttgaagaacaaagatattctttaagaagaagaataatttccctatccccac	789
Db	421	TTTATGTTGAGAACTTAAGATATTTTATTAAGCAAGAGAAATATTTTCCCTTATCCAC	480
Qy	790	tgccttgtaggttctctgaacctgccttgatcgtgtaacgtgtaagatcccggaagccttgccg	849
Db	481	TGCTGTGAGAGTTTCTGTACTCGTGTGGATGGCCGTGAAGATCCCGGAGCCTTGCCG	539
Qy	850	actgccttgtaggtgagcttgtagcctc	875
Db	540	ACTGCTTGTGGTGCGCTTGGCGCTC	565
RESULT	4		
LOCUS	BG532390	665 bp	EST
DEFINITION	602561902F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4699705 5',		
ACCESSION	BG532390		
VERSION	BG532390.1	GI:13523928	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 665)		
AUTHORS	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapub@mail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: CLONETECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.lnl.gov		
	Plate: LNCM1532 row: c column: 02		
	High quality sequence stop: 663.		
FEATURES	Location/Qualifiers		
Source	1..665		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4699705"		
	/clone_lib="NIH MGC 61"		
	/tissue_type="embryonal carcinoma"		
	/lab_host="DH10B (T1 phage-resistant)"		
	/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccttcggcc); Site_2: SfiI (ggccattatggcc);		
	double-stranded cDNA was prepared from cell line RNA. 5'		
	and 3' adaptors were used in cloning as follows: 5'		
	adaptor sequence: 5'-CACGGCCATTAATGCGC-3' and 3' adaptor		
	sequence: 5'-ATTCTTAGAGCCGACGAGCGGCCGACCATG-dt(30)BN-3'		
	(where B = A, C, or G and N = A, C, G, or T). Average		

insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC library."

BASE COUNT 183 a 162 c 181 g 139 t  
ORIGIN

Query Match 60.6%; Score 530.2; DB 11; Length 665;  
Best Local Similarity 91.1%; Pred. No. 2.2e-139;  
Matches 606; Conservative 0; Mismatches 8; Indels 51; Gaps 2;

QY 56 gggagcctacgcttcgctccacacacatgggaatggatgaacaacatcctgcg 115  
DB 1 GGGGGCGCTAGCCTTCCGCTTACAGCCACATGGGGAATGGATTAACAGATCTTGC 60  
QY 116 cctgtacatcggcaactccaagatgccaagagacgcggaacaattgagcaagaag 175  
DB 61 CCTGTACATCGCCACTTCAAGATGCCAGAGACGGGAACATTCAGCAAGAACAGT 120  
QY 176 gacacattctgtctgtccagatagt-ccaagcctatgttgag----- 220  
DB 121 GACACATATTCTGTGTCCAGATAGTCCAGGCGCTATGTTGAGGAGTTAAATACCT 180  
QY 221 -----gacaagacattccaagaagat 244  
DB 181 GTGCATCCGACGCGGATTCACCATCTCAAAACCTGACACAGCATTTCAAAAGATAT 240  
QY 245 taatcatcaccagatgcgcgctccgagtgagagctgcctgtacactgcctgcg 304  
DB 241 TAAATTCATTACGAGATGCGCGCTCCGGTTCAGAGCTGCTTGTACACTGCTGC 300  
QY 305 ggtctcgaagagtgtaactggtgtatcgtacatcgtacacgttaactgacttg 364  
DB 301 GGTCTCCAGAGCGTGACACTGCTGATCCACACATCATGACCTGCTGCTGCTG 360  
QY 365 ggaagatgccttcacacacgctgcgtgcggaagatcctgtgcaaccccaactg 424  
DB 361 GAGAGATGCGCTTCACACCGCTGCTGCTGAGATCTTGTCCACACCCCAACTG 420  
QY 425 ccagagacagctcccaagagtttgaagaagatgagtcacatcgtacgcagtgctga 484  
DB 421 CCAGAGACAGCTCCAGAGATTGAGAAGCATGAGTCCATGATTCGGCAGTGGCTGA 480  
QY 485 ggaagaatctggaagagcccttgcaggaatgcagaagaagccaaacattctgc 544  
DB 481 GGAAGATATGAGAGAGCGCTTTGACAGATCCAGAAAGCAAAACATTCCTGCG 540  
QY 545 tccagaatctcgaagttctggcccttccagaagaagcttaactgaagttctg 604  
DB 541 TCCAGAAATTCGAAGTTCTGGGCTTTTCCAGAAAGCATGTAATGACTGAAATTCG 600  
QY 605 aaatatltgcaaacccgcagagatttaggctggtgtcgtgccaagaagaagaacatag 664  
DB 601 AATATGCAAAACCAACAGAGATTAGGCTGTGTGCCAAAAAAGAAACACATAGAGT 660  
QY 665 ttaag 669  
DB 661 TTAAG 665

RESULT 5  
AV764304 694 bp mRNA EST 19-OCT-2000  
DEFINITION AV764304 MDS Homo sapiens cDNA clone MDSBFH09 5', mRNA sequence.  
ACCESSION AV764304  
VERSION AV764304.1 GI:10922152  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 694)  
AUTHORS Gu, J., Zhao, M., Huang, O., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,  
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng,  
L., Xu, S., Gu, W., Ju, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G.,  
Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.

TITLE Homo sapiens cDNA MDS clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zenguan Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source 1..694

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MDSBFH09"  
/clone\_lib="MDS"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/note="Vector: pTR1pEx2; Site\_1: sf1A; Site\_2: sf1B"

BASE COUNT 183 a 166 c 189 g 154 t 2 others  
ORIGIN

Query Match 59.5%; Score 520.6; DB 10; Length 694;  
Best Local Similarity 88.7%; Pred. No. 1.2e-136;  
Matches 606; Conservative 0; Mismatches 26; Indels 51; Gaps 2;

QY 48 aacagcgcgggcgctagcgttcgccttcacacacatgggaatggatgaacaagatc 107  
DB 1 ACACGGCCGGGCGCTAGCGCTTCGCTTCACGACCATCGGAATGGATGAACAAGATC 60  
QY 108 ctgcccgcgtctacatcggcaactccaagatgccaagagcgcggaacaattggagca 167  
DB 61 CTGCCCGGCTTACATCGGCAACTTCAAGATGCGACAGACCGCAAAATTTGAGCAAG 120  
QY 168 aacaaggtacacatatctgtctgtccacgagatg-ccaagcctatlttgag----- 220  
DB 121 AACCAAGTACACATATTCTGTGTGCCACGATAGTCCAGGCGCTATGTTGAGGAGATT 180  
QY 221 -----gacaagacattcca 236  
DB 181 AATACCTGTGCATCCACAGACGCGATTCCACCATCTCAAAACCTGACAAAGACATTTCA 240  
QY 237 gaaagatataatcattcaccagatgctgcgagtgagagctgcctgtacactgc 296  
DB 241 GAAAGTATTAATTAATTCACAGAGCGCGGCTCCGCGGTGAGAGTGTGCTGTACACG 300  
QY 297 ctgcccgcgtctcgaagagcggtgacactggtgacgataacatagaccgtac 356  
DB 301 CTGGCGGGGCTTCACAGAGCGCTGACACTGCGATCGCATATATACCGTCACTGAC 360  
QY 357 ttggctgggaagatgccttcgacacgctgcgtgcggaagatcctgtgccaaccccaac 416  
DB 361 TTTGGCTGNGAGAGATGCTCCGTCACACCGTGCCTGGGAGATCTCTGTCACACCCAC 420  
QY 417 gtggagctcccgagagcaactccaagatttgaagaagatgagtcacatcagtcag 476  
DB 421 GTGGGCTTCCAGAGACACTCCAGAGATTGAGAAGCTGAGGTTCATCATCTATCCGCG 480  
QY 477 tggctgaagagaataatgagagagcccttgcagagatgcaagaagaagccaaacact 536  
DB 481 TGGCTGAAGGAAGATATGAGAGAGCCCTTTTGCAGGATGAGAGAGAGCAAAACATT 540  
QY 537 ctgcccgtccaggaatctgaaatctcggccttccagaagaagcttaactgactgca 596  
DB 541 CTGGCGGCTTCAGGAATTCAGAGATTCGGGCTTTTTCACAAACAGCTATATGTAACCTGA 600



QY 597 agttctgaataatgcaaacccgcagagtttagctgctgctgcaaaaaaagaaagca 656  
|||||  
Db 601 AGTTCTGAATAATATTCGAACCCACAGAGATTAGGCTGGTCTGTCACAAAAAGGCAACATA 660

QY 657 catagagtttaagatccagtag 679  
| | | | |  
Db 661 GAGTTTGATCAGTAGGATTG 683

RESULT 6  
AM963729 589 bp mRNA EST 01-JUN-2000  
LOCUS EST375802 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.  
DEFINITION AM963729  
VERSION AM963729.1 GI:8153565  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 589)  
Hedge, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.  
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray  
Unpublished (2000)  
JOURNAL Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@tigr.org  
Plate: 192  
Seq primer: Reverse.

FEATURES  
Location/Qualifiers  
source 1..589  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGH"  
/note="Vector: pBluescriptksk"

BASE COUNT 157 a 138 c 162 g 132 t

ORIGIN

Query Match 57.1%; Score 499.6; DB 10; Length 589;  
Best Local Similarity 95.4%; Pred. No. 1e-130;  
Matches 536; Conservative 0; Mismatches 24; Indels 2; Gaps 2;

QY 93 gggatgaacaagatccgctccgctgctgctgctgcaacttcaagaatgcccagagacgc 152  
|||  
Db 1 GGCACCTACGAGATCTCCCGCGGCTGTCATAGCACTTCAAGAGCCAGAGACCG 60

QY 153 gaacaattgagcaagaagaagtgacacatatctgctgctcagatagt-ccagagct 211  
|||||  
Db 61 GACCAATTGACCAAGAAAGGTGACACATATTCGTGTGTCACGATAGTGCCAGGCT 120

QY 212 atgttgagagcaagacattccaagaagaattaaattcaattcagagtgccgagctccg 271  
|||||  
Db 121 ATGTGGAGGAGCAAGACATTTCAAGAAAGTATTAATTCAATTCACGAGTCCGGCTCCG 180

QY 272 cggtagagagctgctgttacaactgctgcccggggtctccagagagctgacactgtgat 331  
|||||  
Db 181 CGGTGAGAGCTGCTTTGACACTGCTGCGGCGGTCTCCAGAGCGTGAACACTGGTAT 240

QY 332 cgcatacatgaccgctgactgctgctgctgagagatgcccgtgcaacgtagcgyc 391  
|||||  
Db 241 CGCATACATCAATGACCGTCTACTGACTTTGGCTGGGAGATGCCCTGCACACCGTGCATGC 300

QY 392 cggagagatccctgtgccaaccccaacgctgagcttccagagagagctccagagattgagaa 451  
|||||  
Db 301 TGGGAGATCTGTGCGCAACCCCAACGTTGGCTTCCAGAGACAGCTCCCGAGAGTTGAGAA 360

QY 452 gcatgagttcatcagatcgcagctgagctgaagaaatcatgagagagcccttgc 511  
|||||  
Db 361 GCATGAGGTCATCATGATGTCGGCAGTGGCTAAGAGAAATATGAGAGAGCCCTTTGCA 420

QY 512 ggaatgcgaagaagcccaaaaacattctgcccgtcccaagaaattcgaagttctggagct 571  
|||||  
Db 421 GGATGCAGAAAGACCCAAAACATTTGCGCGCTCCAGAGATTCGGAAGTTCTGGGCTT 480

QY 572 tctcagaagactgtaatgtagcttctgaaatattgcaaacccgcagagtt-ag 630  
|||||  
Db 481 TCTCAGAAAGACTGTATGTACCTGAATAATTTTGAATAATTCGAACCCCGAGATTAAAG 540

QY 631 gctgtgctgccaagaagaa 652  
|||||  
Db 541 GCTGTCCTTGCACAAAGGAA 562

RESULT 7  
BG674814 950 bp mRNA EST 01-MAY-2001  
LOCUS BG674814  
DEFINITION 602620947F1 NCL\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4746442 5',  
mRNA sequence.  
ACCESSION BG674814  
VERSION BG674814.1 GI:13906210  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 950)  
NIH-MGC http://mgc.ncl.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
Email: cgapds@mail.nih.gov  
Tissue Procurement: James Cleaver, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA  
Sequencing by: Inocyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL0594 row: n column: 11  
High quality sequence stop: 815.

FEATURES  
Location/Qualifiers  
source 1..950  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4746442"  
/clone\_lib="NCL\_CGAP\_Skn3"  
/lab\_host="DH10B (TI phage-resistant)"  
/note="Organ: skin; Vector: pCMV-SPOK6; Site: 1; NotI;  
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.5kb. Library constructed by Life  
Technologies. Note: this is a NCL\_CGAP Library."

BASE COUNT 246 a 225 c 262 g 217 t

ORIGIN

Query Match 52.4%; Score 458.6; DB 11; Length 950;  
Best Local Similarity 81.3%; Pred. No. 4.8e-119;  
Matches 742; Conservative 0; Mismatches 94; Indels 77; Gaps 15;

QY 40 ctgcagaccacgcgcgggagcctagcgtctgccccttcagcccaatgaggaatgagatga 99  
|||||  
Db 1 CTGCGAGACACAGCGGCGCGGCGCTGACGCTTCACCCATCGGGAATGGGATGA 60

QY 100 acaagatccctgcccgcgtgtacatcggcaacttcaaa--gatgcagagacgcggaaca 157  
|||||  
Db 61 ACAAGATCCTCGCCGCGCTGTACATCGGGAACCTTCAATTGATGCGAGAGCGCGAACA 120

QY 158 attgagcaagaagaagtg---acacatatctgtctgtccagata---gtccagagct 211  
|||||

```

Db 121 ATTGACCAATTAAGTAGTACTATATTACTAGTCTGTCCAGATAGATGCCAGTCT 180
QY 212 atgttga----- 219
Db 181 ATGTTGGAGGGAGTTAAGATACCTGTGATCCACAGCGATTCACCGATCTCGAANA 240
QY 220 ----gacacagacattcaaa-gaaagtatttaattcattacagagtcgagccgcgcgt 275
Db 241 CCTTGCAAGACATTTCACACGAAAGTATTAAATTATTCACAGATGCGGCTCCCGAT 300
QY 276 gagaagtccttgtaactatgctgagcggtctccagagagctgacatgtatc-gc 334
Db 301 GAGAGCTGCTGTATACACTGCGCGGCTGTCCAGAGAGGTGACATGATGATCAGC 360
QY 335 ataatcatgacccgtca-ctgaattgtctggagagatgacctgacacc-gtgcgtgc 392
Db 361 ATACATCAATGACCGTATGATGCTTGGCTGAGAGATCCCTGAAACACAGAGCTGCT 420
QY 393 gggagatcctgtgcaaccacacgtgagcttccagagacagctccagagatttgaag 452
Db 421 GGGAGATCTGTGACCAAGCCCAACGTGGGCTTCCAGAGACAGTTCCAGAGATTGAGAG 480
QY 453 catgagttccatcagatcagtcagagtgctga--aggaagatatagagagagcccttgc 510
Db 481 CATGAGGTCCATCAGATGATGCGCAGTGTGACCGGAGACCATATGAGAGAGCCCTTTC 540
QY 511 a-ggatgcgaagaagcccaaaa-cattctgcccgtccagga-atctgaagttctgg 567
Db 541 ATGGATGACGAAGAAAGCAAAATATCATGCTGGCCGTCCGGGACATTCGAGTGTGG 600
QY 568 ccttctcagaagactgtta-tgtaacctgaagttctgaatatatgcaaacccgcagat 626
Db 601 CCTGTGTCAGAAAGACTGTACTGTACTTAAGTTTCTGCCATATGCAAGCCACAGACT 660
QY 627 ttaggtgtgtctgccccaaaaaagaagaacacatagatttaagtaaccagtagtattg 686
Db 661 TTAGGCTGTGCTGCCAAAGTAGAAAGACATAGAGTTAGTATCCAGTAGCATTTTG 720
QY 687 taaactgttttcaatttgaagctgaatatatacgtatgcatgttattatgttgaactaa 746
Db 721 TTAACCTGTGTTCCATTGTAAGCTGATATATATACGATGATGTCATGTCAGAACTAG 780
QY 747 ggaattcttag----caagagaaaaatattcccttacccttaccctgtctggaggtt 803
Db 781 GGATGATGTCGAAGGCAAGGACAAATATTGTCCTTATATACCAATATGTGGAGTTA 840
QY 804 ct-gtaactgcttgatgagctgttaagaatcccgagagccttgcgcacatgcttgg 862
Db 841 CTAGTACCTGCTGATGCTGTAGGATGCCGAGGAGCCTTGCGCACTGCTTGTGG 900
QY 863 tggcttgagctc 875
Db 901 TGGATTGGGGCTC 913

```

```

RESULT 8
LOCUS BG545679 751 bp mRNA EST 04-APR-2001
DEFINITION 602572945F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4701034 5',
ACCESSION BG545679
VERSION BG545679.1 GI:13544344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

```

```

FEATURES
source
1..751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4701034"
/lab_host="DH10B (TI phage-resistant)"
/notes="Organ: Lung; Vector: pMDR-LIB (Clontech); Site: 1;
SfiI (ggcgctcgcc): Site: 2; SfiI (ggcgctcgcc): 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCGGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH-MGC library."
BASE COUNT 189 a 203 c 209 g 150 t
ORIGIN
Query Match 50.9%; Score 445.8; DB: 11; Length 751;
Best Local Similarity 90.7%; Pred. No. 1.9e-115;
Matches 529; Conservative 0; Mismatches 2; Indels 52; Gaps 3;
QY 10 tctctccctgtaacatgacatagtcgctgcgcgaccacagcgccgggagctagcgt 69
Db 2 tctctccctgtaacatgacatagtcgctgcgcgaccacagcgccgggagctagcgt 61
QY 70 cgccttcagccacatggggaatggatgataaagatcctgcgcgctgtacatcggca 129
Db 62 CCCCTTCAGCCACCATGGGGAATGGATGAACAAAGATCTCGCCGCGCTGATACATCGGCA 121
QY 130 actttaaag-atgcagagagcgcggaacaattggaagaagaagaagtacacatatctg 188
Db 122 ACTTAAAGTATGCCAGAGAGCGCGGAACAATTGACCAAGACAAAGTACACATATTCTG 181
QY 189 tctgtccacgatagt-ccaagccctatgttggag----- 220
Db 182 TCTGTCCAGATAGTCCAGAGCGCTTATGTTGGAGGGAAGTTAAATACCTGTGATCCACGA 241
QY 221 -----gacaagacattcaagaagaatataatcaatcac 257
Db 242 GCGGATTCACCATCTCAAAACCTGACAAAGACATTTCAAGAAAGATTAATATCATTCAC 301
QY 258 gattgcgagctcccggtggaagagctgttaacatgctgcgcgaggttccagagagc 317
Db 302 GAGTCCGCGCTCCGCGGAGAGAGCTCTTGATCACTCCGTGCGGCGGCTTCCAGAGAGC 361
QY 318 gtacacatggtatcgcatalacatcagacgctacatgcttggctgtggagagatgcctg 377
Db 362 GTGACACTGTGATGCGATACATCATGACGCTGATGACTTTGGCTGGAGAGATGCCCTG 421
QY 378 caacacgtgctgcgcgagagatcctgtgccaaccacacgtgagcttccagagacagctc 437
Db 422 CACACCGTGTGCTGGGAGATCTGTCGCAACCCCAACGAGGCGTTCACAGACAGCTC 481
QY 438 caggagtttggagagcatgagatcagatagatgagcagtgctggaagaagaatagga 497
Db 482 CAGGAGTTTGAAGCAATGAGGTCCATCATGATGCGCGCTGAGAGGAAGATATATGA 541
QY 498 gagaagccttgcagagatgacgaagaagccaaacatctgg 540
Db 542 GAGAGCCCTTTCAGAGATGCAAGAAAGCCCAAAACATTTCTGG 584

```

Tissue Procurement: Clontech Laboratories, Inc.  
 cDNA Library Preparation: Clontech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.jnl.gov>  
 Plate: LNCM1535 row: 3 column: 11  
 High quality sequence stop: 680.  
 Location/Qualifiers  
 1..751  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4701034"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /notes="Organ: Lung; Vector: pMDR-LIB (Clontech); Site: 1;  
 SfiI (ggcgctcgcc): Site: 2; SfiI (ggcgctcgcc): 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGCGAGCGCGGACATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.9  
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH-MGC library."

RESULT 9  
AA410486 447 bp mRNA EST 18-MAY-1997  
LOCUS  
DEFINITION  
ZV23B12.r1 Soares.NbHMPu\_S1 Homo sapiens cDNA clone IMAGE:754463  
5', mRNA sequence.  
ACCESSION  
AA410486  
VERSION  
AA410486.1 GI:2069654  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 447)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,  
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,  
T., Waterston, R. and Wilson, R.  
Washu-Merck EST Project 1997  
TITLE  
Unpublished (1997)  
JOURNAL  
Contact: Wilson RK  
COMMENT  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28ml3 rev2 EF from Amersham  
High quality sequence stop: 410.  
Location/Qualifiers  
1. 447  
/organism="Homo sapiens"  
/db\_xref="GDB:5977366"  
/db\_xref="taxon:9606"  
/clone="IMAGE:754463"  
/clone\_1lb="Soares.NbHMPu\_S1"  
/tissue\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/note="Organ: mixed (see below); Vector: pT7T3D-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not I;  
Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NbHMP, pregnant uterus  
NbHMP, and fetal heart NbHMP19) were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."  
BASE COUNT 126 a 89 c 115 g 117 t  
ORIGIN  
Query Match 50.5%; Score 442.2; DB 10; Length 447;  
Best Local Similarity 99.3%; Pred. No. 1.8e-114;  
Matches 444; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 181 ATTCTGGCCCTCCAGGAATTCGAACTTCGGCCTTTCACAGAACTGTATGTACC 240  
|||  
Qy 594 tgaagtttcgaaatattgcaaacccgagagtttagctgctccaaaagaaag 653  
|||  
Db 241 TGAAGTTTCGAAATATTCGAAACCCACAGAGTTTAGCGTCTGCCAAAAGAAAG 300  
|||  
Qy 654 caacatagatttaagatccagtagtgattgtaaacctgttttcatttgaagctgaa 713  
|||  
Db 301 CAACATAGAGTTTAAAGATCCAGTAGTAGTATTTGAACTTTTTCATTGGAAGCTGAA 360  
|||  
Qy 714 tatacagtaagcatgttattgttagaactaagagatattcttagcaagagaataatt 773  
|||  
Db 361 TATATAGTACTCATGTTTATGTTGAGACACGAGATATTTTACAGAGAAATATT 420  
|||  
Qy 774 ttcccttattcccaactgctgtgag 800  
|||  
Db 421 TTCCCTTATCCACACGCTGTGAGC 447  
|||  
RESULT 10  
AV711945 502 bp mRNA EST 09-OCT-2000  
LOCUS  
DEFINITION  
AV711945 DCA Homo sapiens cDNA clone DCABH12 5', mRNA sequence.  
ACCESSION  
AV711945  
VERSION  
AV711945.1 GI:10731251  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 502)  
Xu, X., Gu, J., Liu, F., Qu, J., Zhao, M., Li, Y., Huang, Q., Zhou, J.,  
Song, H., Gu, Y., Yang, Y., Gao, G., Xiao, H., Li, N., Qian, B., Gao, X.,  
Cheng, Z., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu,  
G., Cheng, Z., and Han, Z.  
Homo sapiens cDNA clones  
Unpublished (2000)  
CONTACT: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers  
1. 502  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DCABH12"  
/clone\_1lb="DCA"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/note="Vector: pT7Riplex2; Site\_1: sfiI; Site\_2: sfiIb"  
BASE COUNT 142 a 116 c 135 g 109 t  
ORIGIN  
Query Match 44.5%; Score 389.8; DB 10; Length 502;  
Best Local Similarity 95.8%; Pred. No. 1.3e-99;  
Matches 411; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

```

DB 195 ATCATGACCGTACTGACTTTGGCTGGAGGATGCCCTGCACACCGTGGGTGGAGAG 254
QY 399 tctgtgtccaaaccccaagtggtgtccagagacagctccagaggtttgagaagcatgag 458
DB 255 TCGTGTGCCAACCCCAACGTGGGCTTCAGAGACACCTCCAGAGTTTGAAGCATGTAG 314
QY 459 gtccacagatcgcagtcgctgcaagaagaatatgagagagcccttgcagagatgca 518
DB 315 GTTCATCATGATGCGAGTGGCGTGAAGAGAAATATGAGAGAGAGCCCTTTCAGAGATGCA 374
QY 519 gaagaagccaaacaaatcttgcgcgctccagagaatctgaagttctgggaccttccaga 578
DB 375 AAAAAAGCCAAAAACATCTCGCCGCTCCAGGAATCTGAAGTT-TTGGCCTTTTATA 433
QY 579 agatgtatgtacacctgaagttctgaaatatgtcaaccgcgagaggtttagcgtgtgc 638
DB 434 AAACGTAAAGACCCCTGAAGGTTGTGAATATTGCMAAACCCAGAGATTAAAGCTGTGTC 493
QY 639 tgcacaaaa 647
DB 494 TGCACAAAA 502

RESULT 11
BE292273 554 bp mRNA EST 13-JUN-2000
LOCUS 601086078P1 NC1_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500299 5'
DEFINITION mRNA sequence.
ACCESSION BE292273
VERSION BE292273.1 GI:9174552
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 554)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8558 row: k column: 20
High quality sequence stop: 552.
Location/Qualifiers
1..554
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3500299"
/clone_lib="NC1_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 150 a 129 c 158 g 116 t 1 others
ORIGIN

```

```

QY 85 tggggaatggaatgaacaagaatccgtgccgaccttaactgcgaactcaagaatgcca 144
DB 1 TGGGAGTGGGATGAGACCGAGATCTGCCGGGCTGTACATTGGAACCTTCAAGACGCA 60
QY 145 gagacgcggaacaatggaacaagaagtgacacatctctgtgtccagatgt- 203
DB 61 GAGATCCAAACAGTTGACAGGAAACAGAGTACACACTTCTTGTGTGACGATACG 120
QY 204 ccaagcctatgttggagagcaagaacattcaagaagaatlaaatcaatcaagatgc 263
DB 121 CCAGGCCCATGTTGGAGGCAACAGACATTTCAAAGAAAGCATTAATTCATGATGATGC 180
QY 264 cggctccggggtggaagagcgtgctgtgaactgtcctgcggcggtgtccaagaggtgaca 323
DB 181 GCACCTCCAGGTCAGAGCTGCTGTACATTCCTGCGGGGTCTCCAGGAGTGTGACA 240
QY 324 ctggtgatgcatacatcatgacacgtcaactgtgcttggagagatgacctgcacac 383
DB 241 TTGGTATGATGATACATCATGACTGTACACGACTTTGGCTGGGAGATGCCCTTGACACT 300
QY 384 gtgcgtgccggagatcctgtgtccaaaccccaagtcgtccccaagagcgtccagag 443
DB 301 GTTCGTGCGGGAGAGTCTCTGTGCCAACCCCAACTGGCCTTTCAAAGGAGCTGCAGAG 360
QY 444 ttggaagcagatgagtcacatcagratatggcagtgctgcaagaagaatatgagaagac 503
DB 361 TTGTGAAGAAACATGAAGTGCACCATGATCGGCAATGCTGAAGAGACTATGAGAGAAC 420
QY 504 ccttgcagagatgcagaagaagccaaacaaatctcgtg 540
DB 421 CCTTGGCGGATGACGAGAACCAACCAAAATATTCTGG 457

```

```

RESULT 12
AA374753 371 bp mRNA EST 21-APR-1997
LOCUS ESR86937 HSC172 cells 1 Homo sapiens cDNA 5' end similar to similar
DEFINITION to tyrosine phosphatase C1100, mRNA sequence.
ACCESSION AA374753
VERSION AA374753.1 GI:2027073
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 371)
AUTHORS Adams M.D., Kerlavage A.R., Fleischmann R.D., Fudner R.A., Bult
C.J., Lee N.H., Kirkness E.F., Weinstock K.G., Gocayne J.D., White
O., Sutton G., Blake J.A., Brandon R.C., Man Wal C., Clayton R.A.,
Cline T.R., Cotton M.D., Barle-Hughes J., Fine L.D., Fitzgerald
L.M., Fitzhugh W.M., Fitchman J.L., Geoghagen N.S., Glodek A.,
Guelhm C.L., Hanna M.C., Hedblom E., Hinkle P.S.J., Kelley J.M.,
Kelley J.C., Liu L.-I., Marmorek S.M., Merick J.M.,
Moreno-Palauques R.F., McDonald L.A., Nguyen D.T., Fellgrino S.M.,
Phillips C.A., Ryder S.E., Scott J.L., Saudek D.M., Shirley R.,
Small K.V., Spriggs T.A., Uteback T.R., Weidman J.F., Li Y.,
Bednarik D.P., Cao L., Cepeda M.A., Coleman T.A., Collins E.J.,
Dinke D., Feng D.-F., Ferlie A., Fischer C., Hastings G.A., He W.W.,
Hu J.S., Greene J.M., Gruber J., Hudson P., Kim A.K., Kozak D.L.,
Kunsch C., Hungjun J., Li H., Weissner P.S., Olsen H., Raymond L.,
Wei Y.F., Wong J., Xu C., Yu G.L., Ruben S.M., Dillion P.J., Pannon
M.R., Rosen C.A., Haseltine W.A., Fields C., Fraser C.M. and
Venter J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-1/4 (1995)
MEDLINE 96026280
COMMENT Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056

```

Query Match 41.7%; Score 365; DB 10; Length 554;  
 Best Local Similarity 88.8%; Pred. No. 1.4e-92;  
 Matches 406; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

Fax: 3018699423  
 Email: arkerlaw@tigr.org  
 For clone availability, additional sequence and expression  
 information related to this EST, please check the TIGR Human Gene  
 Index (<http://www.tigr.org/cdb/hgi/hgi.html>)  
 Seq primer: M3 Reverse.

# FEATURES

source  
 1. 371  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="ATCC (inhost):179154"  
 /db\_xref="taxon:9606"  
 /clone\_lib="HSCI72 cells I"  
 /cell\_type="fibroblast"  
 /cell\_line="HSCI72 (16PDL)"  
 /dev\_stage="fetal"  
 /note="Organ: lung; Vector: pBluescript SK-; Site\_1: EcoRI  
 ; Site\_2: XhoI"  
 BASE COUNT 93 a 97 c 102 g 78 t 1 others  
 ORIGIN

Query Match 40.8%; Score 357.4; DB 10; Length 371;  
 Best Local Similarity 99.5%; Pred. No. 1.9e-90;  
 Matches 369; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 101 caagatccgcccgcgtgtacatcggcaacttcaagaatgcagagacgcgagacaatt 160  
 |||||||  
 Db 1 CAAGATCCTGCGCCGCTGTACATCGGCAACTTCAAGATGCCAGAGACGGGCAACAT 60  
 QY 161 gagagaagaagaagtgacatatctgtctgtccacgatagt-ccagcgctatgttga 219  
 |||||||  
 Db 61 GAGCAAAACAAAGGTGACACATATCTGTCTGCACAGATAGTCCAGGCGCTATGTTGGA 120  
 QY 220 ggaacaagacattcaagaagatlaattcattcaggatgcggctccgcggtgaga 279  
 |||||||  
 Db 121 GGACAAGACATTTTCAAGAAAGTATTTAAATTTCACAGATGCCGCTCCGCGTGAGA 180  
 QY 280 gctgcctgtacactgtgcgaggggtctccagagagcgtgacactgtgtatcgcataca 339  
 |||||||  
 Db 181 GCTGCTGTGACACTGCTGCGCGGGGCTCCAGAGAGGTGACACTGCTGATCCATATCA 240  
 QY 340 tcatgaccgtcactgaactgtgtcgtgagagatgccctgcacacgcgtgtgcgagagat 399  
 |||||||  
 Db 241 TCATGACCGGTACACTTGGCTGGGAGGATGCTGCACACCGTGTGCTGCTGGAGAT 300  
 QY 400 cctgtgcaaccccaacggtggtcttcagagagagctccagagatttgagaagatgagg 459  
 |||||||  
 Db 301 CCTGTGCCAACCCCAAGCTGGGCTTCAGAGACAGCTCCAGAGATTGAGAGATGAGG 360  
 QY 460 tccatcagat 470  
 |||||||  
 Db 361 TCCATCAGAT 371  
 RESULT 13  
 AA411633 354 bp mRNA EST 17-MAY-1997  
 LOCUS zvl1c08.rl Soares\_NhHMPu\_S1 Homo sapiens cDNA clone IMAGE:753326  
 DEFINITION 5', mRNA sequence.  
 ACCESSION AA411633  
 VERSION AA411633.1 GI:2069430  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wyllie,  
 J., Waterston, R. and Wilson, R.  
 TITLE Washu-Merck EST Project 1997  
 JOURNAL Unpublished (1997)

# COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
 Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 333.

# FEATURES

source  
 1. 354  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5976179"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:753326"  
 /clone\_lib="Soares\_NhHMPu\_S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac  
 (pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte, 2NbH, pregnant uterus  
 NbHPU, and fetal heart NbH19H) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."  
 BASE COUNT 107 a 64 c 79 g 104 t  
 ORIGIN

Query Match 38.4%; Score 336.4; DB 10; Length 354;  
 Best Local Similarity 99.4%; Pred. No. 1.7e-84;  
 Matches 348; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 480 ctgagaagaagaatctggagaagacccttgcagaagtgagaagaagccaagaacattt 539  
 |||||||  
 Db 6 CTGAAGGAGATATGAGAGAGCCCTTGCAGGATG-AGAGAGAGCCAAATCTCTG 64  
 QY 540 gccgtccagaatctctgaatctctggtccttccagaagacttaatgtaactgaagt 599  
 |||||||  
 Db 65 GCCGCTCCAGAAATTTCTAAGTTCTGGGCTTTTCCACAAGACTGTATGTACCTGAGT 124  
 QY 600 ttctgaatatctgaacccgcagagtttaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 659  
 |||||||  
 Db 125 TTCTGAATATTTGCAAAACCCACAGAGTTAGGCTGTGCTGCCAAGAAAGCAACAT 184  
 QY 660 agagtttaagatccagtagatgtgttaactgttttctatttgaagcgtgaatatata 719  
 |||||||  
 Db 185 AGAGTTTAAGTATCCAGTAGTGATGATTTGTAACCTTTTTCATTTGAAGCTGAATATATA 244  
 QY 720 cgtatcatgttctatgttgaagaactaagatattctttagcaagaagaataattccccc 779  
 |||||||  
 Db 245 CGTACTCATGTATTAGTTGAGAACTAAGATATTTTAAAGCAAGAGAAATATTTTCCCC 304  
 QY 780 ttatcccaactgtgtgaggtttctgtacctgcgtgtgagtctgtaag 829  
 |||||||  
 Db 305 TTATCCCACTGCTGTGAGGTTTCTGTACCTCGCTGTGAGATCCCTGTAG 354  
 RESULT 14  
 AA256096 464 bp mRNA EST 13-AUG-1997  
 LOCUS zsz29b10.rl NCI\_GCB1 Homo sapiens cDNA clone IMAGE:686587 5'  
 DEFINITION mRNA sequence.  
 ACCESSION AA256096  
 VERSION AA256096.1 GI:1891635  
 KEYWORDS EST.  
 SOURCE human.



Qy 839 agccttgccgacactgccttctgtgtgtgcttgagcgtc 875  
|||||  
Db 411 Agccttgccgacactgccttctgtgtgtgcttgagcgtc 375

Search completed: February 13, 2002, 01:44:45  
Job time: 4052 sec

**THIS PAGE BLANK (USPTO)**